

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:06:58 ; Search time 67.41 Seconds
(without alignments)
1255.574 Million cell updates/sec

Title: US-09-673-302A-1

Perfect score: 4154
Sequence: 1 GPNICTRGVSSCOQCLAVS.....NNPLYKRNSTFNITRYGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SID5/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SID5/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SID5/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	4119	99.2	788	20 AA49553 Human endothelial
2	4119	99.2	788	20 AA49567 Human antithrombin
3	4119	99.2	788	21 AA492443 Wild type human GP
4	4119	99.2	788	21 AA492443 Human integrin bet
5	4115	99.1	762	21 AA492451 GPIIa variant Arg
6	4115	99.1	762	21 AA492451 Human platelet GPI
7	4114	99.0	762	21 AA492456 GPIIa variant Ser
8	4113	99.0	762	21 AA492448 GPIIa variant Leu
9	4113	99.0	762	21 AA492450 GPIIa variant Ser
10	4112	99.0	762	21 AA492444 Variant human GPII
11	4111	99.0	762	21 AA492452 GPIIa variant Arg

12	4111	99.0	762	21 AA492454 GPIIa variant Pro
13	4111	99.0	762	21 AA492455 GPIIa variant Arg
14	4110	98.9	762	21 AA492449 GPIIa variant Asp
15	4108	98.9	762	21 AA492453 GPIIa variant Cys
16	3844	92.5	787	21 AA492453 Mouse beta-3 integ
17	3768	90.7	718	11 AA492453 Secreted GPIIa su
18	3321	79.9	728	22 AA492453 Novel human extrac
19	3282	79.0	720	18 AA492453 Mouse beta-3 integ
20	2339	56.3	799	17 AA492453 Human integrin bet
21	2156	51.9	788	13 AA492453 B6H. Homo sapiens
22	2156	51.9	788	20 AA492453 Human integrin cel
23	2150	51.8	788	22 AA492453 Beta6 integrin sub
24	1838	44.2	798	22 AA492453 Human polypeptide
25	1831	44.1	798	22 AA492453 Human integrin bet
26	1826	44.0	798	22 AA492453 Porcine CD29 prote
27	1776.5	42.8	805	22 AA492453 Guinea pig integr
28	1676	40.3	577	20 AA492453 B6GP. Cavia porce
29	1636	39.4	557	13 AA492453 Integrin beta-1 ch
30	1608	38.7	963	19 AA492453 Drosophila melanog
31	1537	37.0	845	22 AA492453 Recombinant beta-8
32	1513.5	36.4	769	11 AA492453 LFA-1 beta subunit
33	1513.5	36.4	769	16 AA492453 LFA-1 CD18 subunit
34	1513.5	36.4	769	21 AA492453 Novel human diagno
35	1513.5	36.4	769	22 AA492453 Beta subunit of hu
36	1504.5	36.2	769	13 AA492453 Human LFA-1 beta c
37	1494.5	36.0	769	20 AA492453 Human cancer assoc
38	1375	33.1	699	20 AA492453 Human colon cancer
39	1373.5	33.1	565	21 AA492453 Human CD18 for use
40	1373.5	33.1	565	22 AA492453 Rabbit beta-8 inte
41	1343.5	32.3	676	18 AA492453 Human beta-8 inter
42	1325	31.9	768	13 AA492453 Sequence of a pept
43	1304	31.4	768	13 AA492453
44	1271	30.6	266	14 AA492453
45	1265	30.5	266	14 AA492453

ALIGNMENTS

RESULT 1
AA49553
ID AA49553 standard; Protein: 788 AA.
AC AA49553:
XX
XX
DT 13-JAN-2000 (first entry)
XX
XX
DE Human endothelial membrane glycoprotein IIfa protein sequence.
XX
XX Human: coding sequence polymorphism; vascular pathology gene;
XX
XX polymorphic site; phenotype correlation; forensic; paternity testing;
XX
XX medicine; genetic analysis; vascular disease.
XX
XX OS Homo sapiens.
XX
XX PN W09950454-A2.
XX
XX PD 07-OCT-1999.
XX
XX PE 26-MAR-1999; 99WO-US06473.
XX
XX PR 01-APR-1998; 98US-0054272.
XX
XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
XX PI Lander ES, Daley CO, Cargill M, Ireland JS, Rozen SC;
XX
XX WPI: 1999-620066/53.
XX
XX DR N-PSDB; AA492452.
XX
XX PT Determination of polymorphisms in genes, especially those identifying
XX
XX predisposition to vascular disease

PS Disclosure; Fig 6; 134pp; English.

CC AA232159 to AA232194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AA49550 to AA49573 represent the proteins which correspond
CC to some of the reference alleles.

XX Sequence 788 AA:

Query Match 99.2%; Score 4119; DB 20; Length 788;
Best Local Similarity 99.3%; Pred. No. 6.5e-283;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSPRDLKENLKNCAPESTIEPVS 60
DB 27 gpnicttrgysscgqclavspmcawcsdealplgsprcdlkenllkncapesiefpvs 86
QY 61 ARVLEDRPLSDKSGSSQVTVSPORIALRLRPDDSKNFSIOVROYEDYPVDIYYLMDL 120
DB 87 arvledrplsdksqsgssqvtpqrlalrlrpdsknfsiqrvedypvdiyyimdl 146
QY 121 SYGMKDDLMSTONLGTAKTATQMRKLSNLRIGGAFVDKPVSPYMYISPEALENFCYDM 180
DB 147 sygmkdllmstlgnlgtaklqmrltsnlrlgfatvdkpvsypmyispealeenpcydm 206
QY 181 KTTCLEPMFGYKHLVLTLDQVTRFNEEYKOSVSRNRDAPEGGDAIMQATVCEKIGWRN 240
DB 207 ktcclepmfygkhlvltldqvtrfneevkqsvsrnrndapeggdaimqatvcekiqwrn 266
QY 241 DASHLVEFTDAKTHIALDGRLAGIVQPNDOGCHVSDNHYSASTIMDYPSLGLMTEKLS 300
DB 267 dashllvftdaktahlaldgrlagivqpnogchvsgdnhyssastimdypslglmekls 326
QY 301 OKNINILFENVENVNLXONYSFLIPGTWGVLSMDSSNVLOLIVAVGKISKVLEVR 360
DB 327 oknlnilfenvenvnllyqyselipgtlvglismdssnvlqlivdaygklskvelevr 386
QY 361 DPEBELSLFENATLNNENYIPGLKSCMGKIGDVTSEFSLKAVRGCPQEKESFTIKPYG 420
DB 387 dpeelslfennatlnnenyiplgkscmgkigdvtsfslkavrgcpqekesftikpyg 446
QY 421 FKDSLIVQYTFDCDCACQAOAEPNSHRCNNGTFEGGVCRCGPGWLGSQCECSDEDRP 480
DB 447 fkdsliivqytfddccacqaaepnshrcnngtfeegvcrcgpgwlgscsecsedryp 506
QY 481 SOODECSPREGPVCSORBECLCGQCVCHSSDPKRTGKVCRCDDPSCRYGEMKSGSG 540
DB 507 sqodespregpvcsoorbclcgqcvchssdpkrtgkvcrcddpscrygemkssgsg 566
QY 541 QCSGCDLSDSWTGYTCCTRTDTCMGSNGLICSGRCKCEGSCVCTOPSYGDTCK 600
DB 567 qcsgcdlsdswtgytcctrttdtcmsnglicsgrkcegsccvctopsygdteck 626
QY 601 CPTCPACTFKKCEVCKKFFDRGALHDENTCNRYCDEIESYKELKDTGKDAVNCTYKNE 660
DB 627 cptcpactfkkecvckkfdragalhdentcnrycdeiesyekelkdtgkdaavncitykne 686
QY 661 DDCVAFQYEEDESSGKSLIYVVEEPCPKGPRIYVLSVMAIILIGAILTKMLLT 720
DB 687 ddcvafqyeeedesgksliyvveeppckgpriylvlsvmgaililgailtkmlilt 746
QY 721 IHDRKEFAKFEERARAKMDTANNPYKATSTFTNITYRG 762
DB 747 ihdrkefakfeeararakmdtanmplykatestftnityrgt 788

RESULT 2

AA49567
ID AA49567 standard; Protein; 788 AA.

XX
AC AA49567;

XX
DT 13-JAN-2000 (first entry)

XX
DE Human antithrombin III protein sequence.

XX
KW Human; coding sequence polymorphism; vascular pathology gene;

XX
KW polymorphic site; phenotype correlation; forensic; paternity testing;

XX
XX medicine; genetic analysis; vascular disease.

OS Homo sapiens.

PN W09950454-A2.

PD 07-OCT-1999.

PF 26-MAR-1999; 99WO-US06473.

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

DR WPI: 1999-620066/53.

DR N-PSDB: AA232186.

PT Determination of polymorphisms in genes, especially those identifying

PR predisposition to vascular disease

PS Disclosure; Fig 30; 134pp; English.

CC AA232159 to AA232194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AA49550 to AA49573 represent the proteins which correspond
CC to some of the reference alleles.

XX Sequence 788 AA:

Query Match 99.2%; Score 4119; DB 20; Length 788;
Best Local Similarity 99.3%; Pred. No. 6.5e-283;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSPRDLKENLKNCAPESTIEPVS 60
DB 27 gpnicttrgysscgqclavspmcawcsdealplgsprcdlkenllkncapesiefpvs 86
QY 61 ARVLEDRPLSDKSGSSQVTVSPORIALRLRPDDSKNFSIOVROYEDYPVDIYYLMDL 120
DB 87 arvledrplsdksqsgssqvtpqrlalrlrpdsknfsiqrvedypvdiyyimdl 146
QY 121 SYGMKDDLMSTONLGTAKTATQMRKLSNLRIGGAFVDKPVSPYMYISPEALENFCYDM 180
DB 147 sygmkdllmstlgnlgtaklqmrltsnlrlgfatvdkpvsypmyispealeenpcydm 206
QY 181 KTTCLEPMFGYKHLVLTLDQVTRFNEEYKOSVSRNRDAPEGGDAIMQATVCEKIGWRN 240
DB 207 ktcclepmfygkhlvltldqvtrfneevkqsvsrnrndapeggdaimqatvcekiqwrn 266
QY 241 DASHLVEFTDAKTHIALDGRLAGIVQPNDOGCHVSDNHYSASTIMDYPSLGLMTEKLS 300

```

DB 267 dashllvftctdaktahldgrlagivqpdgchvgsdnhysastctmdypslgltmteklis 326
OY 301 OKNINLIFAVTEENVVNLQYNTSELIPGTTVGLSMDSNVLIQLIDAYGKIRSKVELEVR 360
DB 327 qknlnlffavtenvvnlyqnyseilipgtltvylsmdsenvlqlldaygkirskeleivr 386
OY 361 DLPEELSLSFNATCLNNEVIRPGKSCMGKIGDTVSFSISIAKVRCCPOEKEKSTIRKPVG 420
DB 387 dlpeelstlsfnatclnnevirpglkscmgkigtvsfsisiekvrgcpqekesflkpvgy 446
OY 421 FKDSLIVQTFPCDCACQAQAPNSHRCNNGNTEGCVGRCGPGWLSQCCESEEDRP 480
DB 447 fkdslivqtfpcdcacqaqapnshrcnngntecgvcrcgpgwlsqcceseedryp 506
OY 481 SQODECSPREGQPVCSQRCGECGCVCHSSDFGKITGKCECDPFCVRYKGMCSGHC 540
DB 507 sqgdecspreqpvcsqrgceclcgcvchssdfgkltgkcecdpfcvrykmgcsghg 566
OY 541 QCSGDCDCLDSDMTGYVNCCTTRDTQMSNGILCSGRKCEGSCVCIQPGSYGDTCEK 600
DB 567 qcsgcdclcdsdmtgyvnccttrdtcmsngllcsgrkcegsccvciqpgsygdtcek 626
OY 601 CPTCPDCTFKKECECKKFDGALHDENTCNRRCDEIESYKELKDTGKDAVNTCTYKNE 660
DB 627 cptcpdactfkkececkkfdreymtentcnrycrdelesvkelkdtgkdaavntctyke 686
OY 661 DDCVVFQYEDSSGKSTLYVEEPECPKGPILVVLVSVMAILLIGLAALLIKLIT 720
DB 687 ddcvvfgyedssgkstlyveepecpkypdillvllsvmailllglallikllit 746
OY 721 IHDRKEFAKFEERARAKWDTANPLYEATSTFNITYRG 762
DB 747 ihdrkefakfeerarakwdtanpnykeatsfnityrgt 788

RESULT 3
AA92443
ID AA92443 standard; Protein: 788 AA.
AC AA92443;
XX
XX 10-AUG-2000 (first entry)
DE Wild type human GPIIb, alloantigen P1A1.
XX
XX GPIIb: neurological disease; genotype: phenotype; diagnosis:
XX allantoigen: platelet antigen; P1A1; P1A2: Alzheimer's disease; stroke;
XX neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
XX amyotrophic lateral sclerosis; multiple sclerosis; dementia.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Peptide 1..26
XX FT /Label= signal_peptide
XX FT Protein 27..788
XX FT /Label= mature_protein
XX
XX MO200020634-A1.
XX
XX 13-APR-2000.
XX
XX 01-OCT-1999; 99WO-1B01696.
XX
XX 01-OCT-1998; 98US-0102624.
XX
XX (NOVA-) NOVA MOLECULAR INC.
XX
XX Schapert K;
XX
XX WPI: 2000-303801/26.
XX
XX N-PSDB; AAA09362.

```

```

XX Identifying a subject at risk for a neurological disease comprises
PT determination of genotype or phenotype of GPIIb or GPIIb locus and
PT determining presence of variant GPIIb or GPIIb allele or isoform
XX
XX Disclosure; Fig 3; 55pp; English.
XX
XX Two different forms of GPIIb, alloantigens P1A1 and P1A2 (for Platelet
CC Antigen 1 and 2) have been described and can be distinguished using a
CC monoclonal antibody. The rarer form, P1A2 (see AA09363), has sustained
CC a point mutation at base 192 that causes a nucleotide change from a T to
CC a C and thus a leucine to proline amino acid substitution at residue
CC position 33.
CC Identifying a subject at risk for a neurological disease comprises
CC determining the genotype or phenotype of the GPIIb or GPIIb locus of a
CC subject, and determining the presence of a variant GPIIb or GPIIb
CC allele or isoform, where the presence indicates an increased risk of
CC neurological disease. The method is useful for identifying a risk,
CC diagnosing and treating a neurological disease comprising Alzheimer's
CC disease, neurofibromatosis, Huntington's disease, depression,
CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
CC disease and multi-infarct dementia.
XX
XX Sequence 788 AA:
SQ
Query Match 99.24; Score 4119; DB 21; Length 788;
Best Local Similarity 99.34; Pred. No. 6,5e-283;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 GPNICTTRGVSSCOQCLAVSMCMWCSDEALPLCSPRCDLKENLKNCAPESTEFVSE 60
DB 27 gpnicttrgvsscqclavsmcmwcsdealpysprcdlkenlkncapestefvse 86
OY 61 ARVLEDRPLSDKSGDSQVQVSPORIALRLRPDDSKNFSGVQVEDYVVDIYYVMDL 120
DB 87 arvledrplsdksqdsqvqvsporialrlrpdsknfsigvqvedyvdiyyvmdl 146
OY 121 SYMKDMLSIQNLGKTLATQMRKLTSLNRIGFAGFVDKPVSPYMYISPPALENPCYDM 180
DB 147 symkddmlsiqnlgtklatqmrkltslnrligfagfdkpvspymyisppaleenpcydm 206
OY 181 KTTCLPMFGYKHVLTLDQVTRFNEEVKQKQSVSRNRDAPBEGUVAIKQATYCDCKIGWRN 240
DB 207 ktclpmfgykhvltldqvtrfneevkqkqsvsrnrdapeguvalkqatycdckigwrn 266
OY 241 DASHLIVFTTAKTHIALDGRLAGIVQPDGCHVGSNHRASASTMDYPSLGMTFKLS 300
DB 267 dashllvfttctdaktahldgrlagivqpdgchvgsdnhysastmdypslgltmteklis 326
OY 301 OKNINLIFAVTEENVVNLQYNTSELIPGTTVGLSMDSNVLIQLIDAYGKIRSKVELEVR 360
DB 327 qknlnlffavtenvvnlyqnyseilipgtltvylsmdsenvlqlldaygkirskeleivr 386
OY 361 DLPEELSLSFNATCLNNEVIRPGKSCMGKIGDTVSFSISIAKVRCCPOEKEKSTIRKPVG 420
DB 387 dlpeelstlsfnatclnnevirpglkscmgkigtvsfsisiekvrgcpqekesflkpvgy 446
OY 421 FKDSLIVQTFPCDCACQAQAPNSHRCNNGNTEGCVGRCGPGWLSQCCESEEDRP 480
DB 447 fkdslivqtfpcdcacqaqapnshrcnngntecgvcrcgpgwlsqcceseedryp 506
OY 481 SQODECSPREGQPVCSQRCGECGCVCHSSDFGKITGKCECDPFCVRYKGMCSGHC 540
DB 507 sqgdecspreqpvcsqrgceclcgcvchssdfgkltgkcecdpfcvrykmgcsghg 566
OY 541 QCSGDCDCLDSDMTGYVNCCTTRDTQMSNGILCSGRKCEGSCVCIQPGSYGDTCEK 600
DB 567 qcsgcdclcdsdmtgyvnccttrdtcmsngllcsgrkcegsccvciqpgsygdtcek 626
OY 601 CPTCPDCTFKKECECKKFDGALHDENTCNRRCDEIESYKELKDTGKDAVNTCTYKNE 660
DB 627 cptcpdactfkkececkkfdreymtentcnrycrdelesvkelkdtgkdaavntctyke 686

```

QY 661 DDCVVRFOYEDSSGSKSLVVEEPCPKGPDILVLLSVGAILLGLALLIMKLLIT 720
 Db 687 ddcvvrffyyedsqsksllyvveeepckpddilvllsvmgalllllglaallimkllit 746
 QY 721 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTTYRGT 762
 Db 747 ihdrkefakfeeararakwdtannplykeatsftnttyrgt 788
 RESULT 4
 ID AAY81461 standard; Protein; 788 AA.
 XX AAY81461;
 AC 03-JUL-2000 (first entry)
 DT
 DE Human integrin beta 3.
 XX Integrin beta 3; human endothelial glycoprotein; GP3A; GPIIb;
 KM ITGB3; CD61; platelet glycoprotein 3a; cellular adhesion;
 KM vitronectin receptor; fibronectin receptor; expression inhibition;
 KM antisense therapy; tumour formation; cancer invasion; bleeding disorder;
 KM inflammation.
 XX Homo sapiens.
 OS
 XX US6037116-A.
 PN 14-MAR-2000.
 PD 25-JUN-1999; 99US-0344520.
 XX 25-JUN-1999; 99US-0344520.
 PF (ISIS-) ISIS PHARM INC.
 PR 25-JUN-1999; 99US-0344520.
 XX
 PA Bennett CF, Cowsett LM, Monia BP;
 XX
 PI WPI; 2000-246189/21.
 DR N-PsDB; AAA07028.
 DR
 XX
 PT New antisense compound that inhibits human integrin beta3, useful e.g.
 for treating or preventing infection, inflammation and tumors
 XX
 PS Disclosure; Columns 43-48; 33pp; English.
 XX
 CC This sequence represents human integrin beta 3. Integrins constitute
 one of four classes of cellular adhesion molecules, and play an
 CC important role in cell migration, cell anchorage to substrates and
 CC cytoadhesion signalling pathways. They are heterodimeric
 CC cation-dependent membrane glycoproteins composed of an alpha and beta
 CC subunit. Integrin beta 3 (also known as human endothelial glycoprotein,
 CC GP3A, GPIIb, ITGB3, CD61 and platelet glycoprotein 3a) is the common
 CC beta subunit partner of the members of the beta-3 subfamily of integrins.
 CC This family consists of the vitronectin receptor (alpha-v-beta-3) and the
 CC fibronectin receptor (alpha-iiB-beta-3). Cells expressing this class of
 CC integrin can adhere to various matrix proteins and participate in
 CC various cytoadhesion-driven cellular responses. Integrin beta 3 is
 CC implicated in conditions such as vascular restenosis, excessive bone
 CC resorption, angiogenesis (in melanoma), tumour invasion, platelet
 CC aggregation and glanzmann's thrombasthenia. The invention relates to
 CC antisense oligonucleotides targeted to the human integrin beta 3 gene,
 CC which inhibit its expression. A series of oligonucleotides
 CC (AAA07035-AA07074) were designed to target different regions of the human
 CC integrin beta 3 RNA, and were analysed for their effect on integrin beta
 CC 3 mRNA levels by quantitative real-time PCR. The oligonucleotides of the
 CC invention are useful for diagnosis, prevention and treatment of
 CC conditions associated with integrin beta 3 expression, such as tumour
 CC formation, inflammation, infections and the diseases mentioned above.
 CC
 XX Sequence 788 AA;
 SQ

Query Match 99.2%; Score 4119; DB 21; Length 788;
 Best Local Similarity 99.3%; Fred. No. 6.5e-285;
 Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 GPNICTTRVSSCOOCLAWSPMCASDEALPLGSPKCDLKENLKDCAPESTIEFPVSE 60
 Db 27 gpnicttrvssccqqlavspmcawcdealpigsprcdlkenlkdncapesiefpyse 86
 QY 61 ARVLEDRPLSDKSGSDSSQVTVSPQRIALRLRPDSKFNFSI QVRQVEDYEVVDIYIMDL 120
 Db 87 arvledrplsdkgsgdsqvtvspqriralrlrpdskfnfsiqvrqvedyevvdiylmld 146
 QY 121 SYSMKDDLSIQNLGTAKTATQKRKITSULRIGFGAFVYKPKVPYRMYISPPALENPGCDM 180
 Db 147 sysmkddlswsiqnlgtaklatqmrkltslrlyfgafvdkpvsrmyisppalempcydm 206
 QY 181 KTTCLPMFGYKHHVLTLPDQVTFNFEVEVKOSVSRNRDLAPEGGUAIMQATVCDERIGWRN 240
 Db 207 ktclpmyfgykhvltldqvtfnfeevkksvsnrmdpeggidaimgalvcdekigwrn 266
 QY 241 DASHLLVFTTDAKTHIALDGLAGIVQPNDCQCHVGSNNHYSASTMDYPSLGIMTEKLS 300
 Db 267 dashllvfttdakthialdglagivqpnndqchvgsnnhysastmdypslgimtekls 326
 QY 301 OKNINLIRAVTENYVNLQNTSELIPRTTVGLSMDSSNNVQLIYDAVGRKRSVELEVR 360
 Db 327 oknlnliravtenvnlqntseliprttvglmsdssnnvqllydaygrkrsvelevr 386
 QY 361 DLPEELSLSPNATCLNNEVIFGLKSCMGLKIGDYNSEFIEKAVGCPDEKESSTIRPVG 420
 Db 387 dlpeelstspnatclnnevifglkscmglkigdvnsefiekavgcpekeksstirpvg 446
 QY 421 FKDSLIVVYTPDCCACQAQAEAPNSHRCNNGNTPFEGCVRCRCRWLTASQCESEEDYRP 480
 Db 447 fkdsllvvyltpddcacqaqaepnshrcnngntpfecvrcrcrwltagseceedyrp 506
 QY 481 SQODECSPREGQPVCSQSGECCLGCVCHSSDPFKITGKYCEGDFSCVRRKGMCSGHG 540
 Db 507 sqodecspregpvcsqsgecclgcvcvchssdfkitgkycegdfscvrrkgmcsghg 566
 QY 541 QCSGDCLCDSDWMTGYCNCRTTRDTGCMSSNGLLCSGRKCECSCVCIOPGSGYDTCER 600
 Db 567 qcsgcdclcdsdwmtgycncrttrdtgcmssnllcsgrkcecscvciopgsgydtcer 626
 QY 601 CPTCPDACTFRKKEVECKEKKPRGALHDETCNRCRDEFSVRLTKTGDAVNCYTKNE 660
 Db 627 cptcpdactfrkkeveckkprgalmhdeTCNRCRDEFSVRLTKTGDAVNCYTKNE 686
 QY 661 DDCVVRFOYEDSSGSKSLVVEEPCPKGPDILVLLSVGAILLGLALLIMKLLIT 720
 Db 687 ddcvvrffyyedsqsksllyvveeepckpddilvllsvmgalllllglaallimkllit 746
 QY 721 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTTYRGT 762
 Db 747 ihdrkefakfeeararakwdtannplykeatsftnttyrgt 788
 RESULT 5
 ID AAY92451 standard; Protein; 762 AA.
 XX AAY92451;
 AC 10-AUG-2000 (first entry)
 DT
 DE GPIIb variant Arg214Gln.
 XX
 DE GPIIb; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KM allantoic acid; platelet antigen; PLAI; PLAI2; Alzheimer's disease; stroke;
 KM neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
 KM anyotrophic lateral sclerosis; multiple sclerosis; dementia.
 KW

OS Homo sapiens.
 XX
 FN WO200020634-A1.
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-IB01696.
 XX
 PR 01-OCT-1998; 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K;
 DR WPI; 2000-303801/26.
 XX
 PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIIa or GPIIb locus and
 PT determining presence of variant GPIIIa or GPIIb allele or isoform
 XX
 PS
 XX
 XX Disclosure; Page -: 55pp; English.
 CC Known polymorphisms in GPIIIa that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GPIIIa sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification. It was created
 CC from the wild type human GPIIIa sequence which appears in Figure 3.
 XX
 XX Sequence 762 AA;
 XX

Query Match	99.1%	Score 4115:	DB 21:	Length 762:
Best Local Similarity	99.2%	Pred. No.1.2e-282:		
Matches 756:	Conservative 1:	Mismatches 5:	Indels 0:	Gaps 0:
QY 1	GNPITTRGVSSCQOCLVAVSPMCANCSDEALVGSPRCDLKENLKDNCAPESIEFPVSE	60		
DB 1	gnpicttrgvsscgcqlavspmcawcsdealvgprcdlkenllkdncapesielfpvse	60		
QY 61	ARVLEDRPLSKGSSDSSQVTVQVSPQRALRLRRPDSKKNFSIQVQVQVEYPRVDIYYLML	120		
DB 61	arvledrplsdgsgdsasvqcvspqraltlrtpdsknfsiqvqvveyprdiyyiml	120		
QY 121	SYSMKDDLSIONLSTKATLQMKRLKTSNLRIGFAFVDRVPVPMYITSPREALENPCYDM	180		
DB 121	ysymkddlsionlgtklatkqmrkltsnrlrgfatvdkrvpsymyisprealennpcydm	180		
QY 181	KTTCLPRMGKYKIVLLTDQVTRFRNEBYKKOSVSRNRDAEGGFDAIMQATVCDERIGMKN	240		
DB 181	kttclpmrgkykhnvlltldqvtrfrneevkkqvsqnrirdapggfdalimgatvcdckigwtn	240		
QY 241	DASHLLEVTDAKTHIALDGLAGLGIYOPRDGCGCHNSDNNHYSASTMDVPSGLIMTERKS	300		
DB 241	dashllvftdckthialdgrlgiyivprndggcghnsdnhysastmdvpsglimterks	300		
QY 301	QKNINLIEAVTENVYVNLKONYSELIPGTTVGVLSMDSSNVLDLIDAYGKINSKYLEVR	360		
DB 301	qkninliefavtenvynlyngyselipgttvgvlsmdssnvlqllivdaygkiskylevr	360		
QY 361	DLPEELISLFNFNMCCLNNEVYIPGLSKSGMGLKIDGYVFSLEAKVRGCPQKEKSEFTTKPVG	420		
DB 361	dlpeelislfnfncclnnevypglkskgmglkldgyvfsleakvrgcpqekesftlkpyg	420		

Qy	421	FKDLSITVOVFEDDCCACAOAEPHSHCNNGNFTFEGCVCRCRGMLTSGCECSEEDYR	480
Db	421	fKdSlIvqVtltdccacagadepnshrcmngnlftecgvcccgjwJygsqcecedyR	480
Qy	481	SQOEDCSPREGQOPVCSORGEELCCQCVCHSSDFEKLTKGKCECDDEFSVRYKGBMCSGC	540
Db	481	sqdecspreegqpyvcsqrgelccqvcchssdfgkltgkycecddefsvrykgemcsgh	540
Qy	541	QCSGCDCLCDSDMWTGYCCNCTRTTDTDCMSNGLLCSGRGKCEGCSVC10PGSVGDTCEK	600
Db	541	qcsgcdclcdsdwgyccnctrttdtcdmsngllcsgrykcegsvc1qpsvgdtcek	600
Qy	601	CPTCPDACTFEKKECBEVECKEDRGALHDBNTCNRYKCRPEIESVKFLKJTGADAVNCTYKNE	660
Db	601	cptcpdactfkekevceckf4repymentcnrycrdalesvkflkjtgdavnclykne	660
Qy	661	DDCVVRQYVYEDSSGKSLTYVEEPCEPKGDIIVVLLSWGATLLGLAALLIWKLLIT	720
Db	661	ddcvvrtyayyedsagkslltyveepecpkgdilivvllswgailllglaalllwkllit	720
Qy	721	IHDRKEFAKFEERBARAKMDANPNLVKEATSTFTNTTYRGT	762
Db	721	IhdrkefakfeeerarawdampLVkEaatstftntlyrgt	762

RESULT	6
AARS1500	
ID	AARS1500 standard; Protein; 788 AA.
XX	
AC	
XX	AARS1500;
XX	
DT	17-NOV-1994 (first entry)
XX	
DE	Human platelet GPIIa surface antigen YUK-a.
XX	
KW	platelet surface antigen; glycoprotein GPIIa; YUK antigen; probe;
RN	detection; thrombocyte; YUK-a allele.
XX	
OS	Homo sapiens.
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 169
FT	/note= "Yuk-a and Yuk-b differ at position 169 with
FT	Gln (corresp. to CAA codon) in Yuk-a but
FT	Arg (corresp. to CGA codon) in Yuk-b"
FT	
FM	
PM	JR06078771-A.
XX	
PD	22-MAR-1994.
XX	
PF	06-JUL-1991; 91JP-0192554.
XX	
PR	06-JUL-1991; 91JP-0192554.
XX	
PA	(JUJI/) JUJI T.
XX	
DR	WPI; 1994-131276/16.
DR	N-PSDB; AAO62199.
XX	
PT	The detection of platelet surface antigen determinant - using
PT	oligo-nucleotide probes for YUK a and YUK b
XX	
PS	Disclosure; Fig 1-5 (Page 8-12); 12pp; Japanese.
XX	
CC	Primers (3) and (4) (AAO62195 and AAO62196, respectively) were used to
CC	amplify a region of the YUK gene from platelet cDNA derived from
CC	humans of known YUK antigen type. The 1mer probes AAO62193 and
CC	AAO62194, specific for the YUK-a and YUK-b alleles respectively, were
CC	tested for hybridisation to the cDNA. The YUK-a probe hybridised to
CC	cDNA from a YUK a/a homozygote and from YUK a/b heterozygotes only,
CC	while the YUK-b probe hybridised to cDNA from the heterozygotes and
CC	from YUK b/b homozygotes only. AARS1500 is the sequence for the human
CC	YUK-a protein.

XX	Sequence	788	AA;
SQ			

Query Match	99.18;	Score 4115;	DB 15;	Length 788;
Best Local Similarity	99.28;	Pred. No. 1.2e-282;		
Matches 756; Conservative	1;	Mismatches 5;	Indels 0;	Gaps 0

QY	1	GNIGICTRVSVCOCGLAVSPMCACSPDEALPLICSPRDLKENTILKONCAPESTIEFPVSE	60
Db	27	gnicttrvsscgcqlavspmcacspdealpisprrcdkentilknccapesiefpvse	86
QY	61	AAVLEDRPLSDKSGSDSSQVTVSPQRTALRLRPDDSKNFESIQRQVEDYPVDTIYTLMDL	120
Db	87	arvledrplsdkgsgsdssqvqvspqrtrialrlrppdsknfisqrvqvdytvdvdiylmcl	146
QY	121	SYSMKDDLMSTIONLGTATQMRKITSLRLIGFAPYDKPVSPLYMTISPEALENPCYDM	180
Db	147	sysmkddlmsionlgtatqmrkittslrlirfigafvdkpvsplymtispealenpcydm	206
QY	181	KTTCTCPMGKYHVLTLTQVTRFPMDEYKKSQSVSNRRAPAEQGFALIMQAVTCEKIGMRN	240
Db	207	kttcpcpmgkyhvlvtlqovtrfpmdeykkssvsnrrapaeqgfalimqavtceklgmrn	266
QY	241	DASHLVEFTTDAKTHIALDGLAGIIVQPNDOQCHVGSDNMHSASTMDYPSLGLMTEKLS	300
Db	267	dashllvfttdakthialdglagivqpnqdqchvgsdnmhsastmdypslglmteklis	326
QY	301	QKNINLIRAVIENVNLYQNTSELPIGTIVGLSMDSNNVLQILVAYGKIRSKVELEVR	360
Db	327	qkninliravenvnnyqnselipgtivglsmdsnnvlqilvdavgkirskevelevr	386
QY	361	DLPEELSLSPNATCLNNBEVIFGLKSCGKIGTIVPSIIEAKYRGCPQEKESFTTKPVG	420
Db	387	dlpeelsslspnatclnnbevifglkscmgltgtdvtsiieakvrgcpqekesftlkpvq	446
QY	421	FKDSLIIYOVTEDCDACQAOAEPNHSRCONNGNTFEGCVCRCGPGWGLSQCECSEEDYRP	480
Db	447	fksdliiyovtfdcdaqaqagapnshrcnngntfegcvcrcgpgwglsgcscseedyrp	506
QY	461	SOQDECSREBQPPVCSQRBETLGCQCYCHSSDRKLTGKICECDDESCVARKGEMCSGHG	540
Db	507	sqgdecspreqppvcsqrgetlqgcvcchssdrltgkycecddfscvrykgemcsghg	566
QY	541	QCSGDCDLCDSDMFGYVCNCTPTPTDMQSSNGLICSGRGKCEGSCVCIOPSGYGTCKR	600
Db	567	qcsgcdclcdsdmfgycncttptptdmqssngllicsgrgkcegsccvciopsgygtckr	626
QY	601	CPTCPDACTFFKKECVECKEDRGALHDENFCNRYCRBELIESVRELKDTGDAVANCYKNE	660
Db	627	cptcpdactffkkecvckldregymtencnzyrcdelesvkeltdgtadvancnykne	686
QY	661	DDCVARRQYYWDSGSKILIVVEPECPKRPDLIVLISMGAILLIGLAALLIMKLLIT	720
Db	687	ddcvarrqyywdsqskilivvepecpkqpdliivlismgailllgliaaillwklilit	746
QY	721	IHDRKEFAKFEESERAKMPTANNPLIKELASTFTNTIYGT	762
Db	747	ihdrkefakfeeserakmptannplykeastsftntilygt	788
RESULT 7			
AAV92456			
ID AAV92456 standard; Protein: 762 AA.			
XX			
AC	AAV92456;		
XX			
DT	10-AUG-2000 (first entry)		
XX			
DE	GP11a variant Ser752pro.		
XX			
GP11a; GP11b; neurological disease; genotype; phenotype; diagnosis; alloantigen; platelet antigen; PLA1; PLA2; Alzheimers disease; stroke;			

KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
xy

OS Homo sapiens.

PN WO200020634-A1
XX

PD 13-APR-2000.
XY

PE 01-OCT-1999; 99WO-IB01696.
XX

PK 01-001-1998; 9805-0102624.
XX

FA (NOVA-) NOVA MOLECULAR INC.
XX

F1 Schiaparelli n,
XX

[illegible]

PT determination of genotype of GPIIb or GPIIb locus an

PS Disclosure; Page -; 55pp; English.

CC Known polymorphisms in GPIIa that may be determined to be variants
CC using the methods of the invention include AAY92447-56, which are variant
CC GPIIa sequences derived from the mature protein.
CC Identifying a subject at risk for a disease associated with a variant

identifying a subject at risk for a neurological disease comprises determining the genotype or phenotype of the GRP114 or GRP130 locus of a subject, and determining the presence of a variant GRP114 or GRP130 allele or isoform, where the presence indicates an increased risk of neurological disease. The method is useful for identifying a risk, diagnosing and treating a neurological disease comprising Alzheimer's disease, neurofibromatosis, Huntington's disease, depression, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease and multi-infarct dementia.

CC from the wild type human GPIIa sequence which appears in Figure 3.

sq Sequence 762 AA;

Query Match	99.08;	Score 4114;	DB 21;	Length 762;
Best Local Similarity	99.28;	Pred. No. 1.4e-282;		
Matches 756; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;

QY 1 GPNICTTRVSSCCQCLAVSPMCAMCSDEALPLGSPCKDKENILKDNCAPESTIEFPVSE 60
Db 1 gpnicttrgvssccgclavspmcawcsdealpysprcdlkenllkdncapestiefpyse 60

QY 61 ARVLEDRPLSDKSGSSQVTQVSPQRIALRLRPDDSKNFSIQVRQVEDYPVDIYYLMDL 12

61 arvledrplsdkgsgdssqvtqvspqrjalrlrrpddsknfslqvrqvedypvdilylmdl 12

121 SYSMKDDLSIQNLGIKLAIQMKRLISNLRIGFGAFVDRKVSPLYMYISPPALENPCYDM 18

121 sysmkaawsiqulqkrlacqmrklcsnrrlgigalvdkpvsppymylsppealenpcyam 180

87 101 K1 CCFM91NHVLI1DQ1KFNNEVAKQSVSRNKRDAFEGGI DAIHQAI VCDEN IGWRN 240

[illegible][illegible]

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

On: 11/11/2019 11:11:11 AM

100

Db 361 dlpeelslfnatclnnevlpglkscmglkigtvtsfisleakvrgcpqekesftlkpvq 420
 Qy 421 FKDSLIVQVTPDCACAOAEPNSHRCNNGNFGFECCVRCRCGPMGLSGOCESEEDYR 480
 Db 421 fkdslivqvtldcdacgaqaepnshrcnmngnfgfecgvcrcgpmglsgocseedytr 480
 Qy 481 SQODECSPREGQPVCSQRCGECGCVCHSSDFGKITGKYCECDPFCVRYKGMCSGHC 540
 Db 481 sqgdecspregqpvcsqrgceclcgcvchssdfgkitgkyceddfacvrykgemcsghg 540
 Qy 541 QCSGDCDCLDSWMTGYCNCCTRTDTCMSSNGLCSGRKCEGSCVCYIOGSGDTCEK 600
 Db 541 qcsgcdclcdsdwmtgycncctrtctdcmssngllcsgrkcecgscvcyioqsgydtcek 600
 Qy 601 CPTCDACCTFKKECECKKFRGALHDENTCNRCRDIESEVKELKDTGKDAVNCYKNE 660
 Db 601 cptcdactfkkececkkfdrdeymtenctnrcyrcdeiesvkelkdtgkdaavncykne 660
 Qy 661 DDCVVRFOYEDSSGKSILYVEEPECPCGPDILVLLSVGAILLIGLAALLIMKLLIT 720
 Db 661 ddcvvrfoyledssgksilyyveepecpkipdillvllsvmgailllglaallimkllit 720
 Qy 721 IHDRKEFAKFEERARAKWDTANNPLLYKEATSTFTNTTYRGT 762
 Db 721 ihdrkefakfeerarakwdtannplykeatstftntlyrgt 762

RESULT 8

AAV92448
ID AAV92448 standard; Protein: 762 AA.

AAV92448;

10-AUG-2000 (first entry)

GP11a variant Leu117Trp.

KW GP11a; GP11b; neurological disease; genotype; phenotype; diagnosis;
 KW allonigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

PF 01-OCT-1999; 99WO-1B01696.

PR 01-OCT-1998; 98US-0102624.

PA (NOVA-) NOVA MOLECULAR INC.

PI Schappert K;

DR WPI: 2000-303801/26.

PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GP11a or GP11b locus and
 PT determining presence of variant GP11a or GP11b allele or isoform
 XX
 XX
 PS Disclosure: Page -: 55pp; English.

CC Known polymorphisms in GP11a that may be determined to be variants
 CC using the methods of the invention include AAV92447-56, which are variant
 CC GP11a sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GP11a or GP11b locus of a
 CC subject, and determining the presence of a variant GP11a or GP11b
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's

CC disease, neurofibromatosis, Huntingdon's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification. It was created
 CC from the wild type human GP11a sequence which appears in Figure 3.

XX Sequence 762 AA:

Query Match 99.0%; Score 4113; DB 21; Length 762;

Best Local Similarity 99.2%; Pred. No. 1,66-282;

Matches 736; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPGSPRCDEKENILKDNCAPIESIEPVSE 60
 Db 1 gpnictrgvsscgqclavspmcawcsdealpisprrcklenilndncapietpvse 60
 Qy 61 ARVIEDRPLSDKSGSSQVTOYSPQRIALRPPDSKNFSTVROVEDYVPVITLMDL 120
 Db 61 arviedrplsdksqssqvtqyspqrlalrppdsknfsivrvedyvpvilyymdl 120
 Qy 121 SYSWKDLMSTIONLGTKLATQMRKITSNLRIFGAFVDPKVPSPYMYISPEALENFCYDM 180
 Db 121 syswkdlmstionlgtklatqmrkitsnlrigafvdkvpvpymyispealenpcydm 180
 Qy 181 KTTCLPMFGYKHVLLTLDQVTRNEEYKOSVSARNRDAPEGHDAIMQATVCOEKIGWRN 240
 Db 181 ktclpmyfgykhvlltldqvtneeykksvsarnrdapeghdaimgatvcoekigwrn 240
 Qy 241 DASHLVFTTDDATHTALDGRLAGIYQPNDCOCHVSDNHYSASTTMDVPSGLMTEKLS 300
 Db 241 dashlvfttdaktthalldgrlagivpndqchvsgdnhyasatmdvpsglmteks 300
 Qy 301 QKNINLFAVTENVNLQYVSELIPTGTGVLSMDSNVLCVDAVGRIRSKVELEVR 360
 Db 301 qkninlfahtenvnlyqyveliptgtgvlsmdsnvlgivdagyrskvelevr 360
 Qy 361 DLPEELSLSFNATCLNNEVLPGLKSCMGLKIGTVSEFSIEAKVRGCPQEKESFTTKPVG 420
 Db 361 dlpeelslfnatclnnevlpglkscmglkigtvtsfisleakvrgcpqekesftlkpvq 420
 Qy 421 FKDSLIVQVTPDCACAOAEPNSHRCNNGNFGFECCVRCRCGPMGLSGOCESEEDYR 480
 Db 421 fkdslivqvtldcdacgaqaepnshrcnmngnfgfecgvcrcgpmglsgocseedytr 480
 Qy 481 SQODECSPREGQPVCSQRCGECGCVCHSSDFGKITGKYCECDPFCVRYKGMCSGHC 540
 Db 481 sqgdecspregqpvcsqrgceclcgcvchssdfgkitgkyceddfacvrykgemcsghg 540
 Qy 541 QCSGDCDCLDSWMTGYCNCCTRTDTCMSSNGLCSGRKCEGSCVCYIOGSGDTCEK 600
 Db 541 qcsgcdclcdsdwmtgycncctrtctdcmssngllcsgrkcecgscvcyioqsgydtcek 600
 Qy 601 CPTCDACCTFKKECECKKFRGALHDENTCNRCRDIESEVKELKDTGKDAVNCYKNE 660
 Db 601 cptcdactfkkececkkfdrdeymtenctnrcyrcdeiesvkelkdtgkdaavncykne 660
 Qy 661 DDCVVRFOYEDSSGKSILYVEEPECPCGPDILVLLSVGAILLIGLAALLIMKLLIT 720
 Db 661 ddcvvrfoyledssgksilyyveepecpkipdillvllsvmgailllglaallimkllit 720
 Qy 721 IHDRKEFAKFEERARAKWDTANNPLLYKEATSTFTNTTYRGT 762
 Db 721 ihdrkefakfeerarakwdtannplykeatstftntlyrgt 762

RESULT 9

AAV92450
ID AAV92450 standard; Protein: 762 AA.

AAV92450;

10-AUG-2000 (first entry)

XX GP11a variant Ser162Leu.
 DE
 XX
 KW GP11a; GP11b; neurological disease; genotype; phenotype; diagnosis;
 KW allantoic acid; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX
 OS Homo sapiens.
 XX
 PN WO200020634-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-IB01696.
 XX
 PR 01-OCT-1998; 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K;
 DR WPI; 2000-303801/26.
 XX
 PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GP11a or GP11b locus and
 PT determining presence of variant GP11a or GP11b allele or isoform
 XX
 PS Disclosure; Page -: 55pp; English.
 XX
 CC Known polymorphisms in GP11a that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GP11a sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GP11a or GP11b locus of a
 CC subject, and determining the presence of a variant GP11a or GP11b
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GP11a sequence which appears in Figure 3.
 CC
 SO Sequence 762 AA:

Query Match 99.0%; Score 4113; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 1.6e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPRICTRGVSSQOCLAVSPKAMCSDEALPLGSPRCLKKNLKNKNCAPSTIEFPVSE 60
 DB 1 gprictirvsscqgclavspmcawcsdealpigsprclkenllkncapesielfpvse 60
 QY 61 ARVLEDRPSPDGSGDSQVSPORIALRLRPDDSKNFSIQVQVEDYPDVIYIMDL 120
 DB 61 arvledrpisdgsgdsqvtvsporiarlrpddsknfisqvgvvelypvdiylmdl 120
 QY 121 SYSMKDDLMSIONLGTIKLATQMRKLTSLNIGFAGFVDPKSPVYVISPPEALENPPCYDM 180
 DB 121 sysmkddlmsionlgtiklatqmrkltslnigfagfvdvdkvpiipymyisppaalencydm 180
 QY 181 KTTCLPMGKXKHLVLTQVTRFNEFEVKQSSRRNDAPEGFDALIMQATVDEKIGWRN 240
 DB 181 ktclcpmgykhvltlqvtvrfneevkqssrrndapegfdalimqatvdekigrwn 240
 QY 241 DASHLVFTTDAKTHIALDGLAGIYQPNDGCHVSDHNYASTTMDYPSGLTETKIS 300
 DB 241 dshllvfttdakthialdglagiyqpn dgchvshdnyasttmdypsgltetkls 300
 QY 301 QRNINLIFAVTENVNLYQNYSELIPGTTGVLSMDSNVQLIVDAYGKIRSKVELEVR 360

DB 301 qkninlifavtenvnlyqnyse lipgttgvlsmdssnvqlivdaygkirkvelevr 360
 QY 361 DLPEELSLSFNATCLNNEVYIPGLKSCMGKLGDIYVSFTEKKNVGCPEKKEKSTIRPVG 420
 DB 361 dlpeelstsfnatclnnevypglkscomglikgdtvstfiekvypcekeksftlirpv 420
 QY 421 FRDSLIVQTPDCACAOAEPNHRCNMGNFEEGCVKCGGKGMGSOCESEEDYRP 480
 DB 421 frdslivqtpdcacaaopnhrcnmgnfegcvkcggkmgsoceceedyrp 480
 QY 481 SQODECSPPREGQPVCSQREGLCGQCVCHSSDFKITKCYCEPDFSCVRRYKEMCSGHS 540
 DB 481 sqodecsppregqpvcsqrgelcgqcvchssdfkitkcycepdfscvrrykgemcsgh 540
 QY 541 QCSGCDCLDSDMWGYTCNCTTRDTCSSNGLLCSGKCGKCGCYCICPGSGDPICEK 600
 DB 541 qcsgcdcldsdmwgytcncttrdtcssngllcsgrkcgkcgcyvicpgsygdlcek 600
 QY 601 CPTCPDACPFRKKECECKKFDKRGALHDENTCNRYCROEIESVKELKDPKDAVACTYKNE 660
 DB 601 cptcpdactfrkceveckkfdrepmtenctnrycdeiesvkelkdpkdaavactykne 660
 QY 661 DDCVVRFOYEDSSGKSLIYVEEPECKGPDIIIVLISVGAALLIGLAALLIMKLLIT 720
 DB 661 ddcvvrfoyedssgksliyyveepecpdpdiilvllsvgaalllglallimkllit 720
 QY 721 IHDKEKFAKFEEDERAKWDTANNPVLYKEATSTNTNTYKGT 762
 DB 721 ihdkefakfeederakwtdanmplykeatsntntitygt 762

RESULT 10
 AAY92444
 ID AAY92444 standard; Protein; 788 AA.
 XX
 AC AAY92444;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Variant human GP11a, allantoic acid.
 XX
 KW GP11a; GP11b; neurological disease; genotype; phenotype; diagnosis;
 KW allantoic acid; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..26
 FT Protein /label= signal-peptide
 FT Msc-difference 27..788
 FT /label= mature-protein
 FT /label= I33P
 FT /note= "mutation caused by T192C substitution in
 the coding sequence"
 FN WO200020634-A1.
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-IB01696.
 XX
 PR 01-OCT-1998; 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K;
 DR WPI; 2000-303801/26.
 DR N-PSDB; AAA09363.

XX Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIb or GPIIb locus and
 XX determining presence of variant GPIIa or GPIIb allele or isoform
 PS Claim 16; Fig 4; 55pp; English.

CC Two different forms of GPIIa, allantoicins P1A1 and P1A2 (for Platelet
 CC Antigen 1 and 2) have been described and can be distinguished using a
 CC monoclonal antibody. The rarer form, P1A2 has sustained a point mutation
 CC at base 192 that causes a nucleotide change from a T to a C and thus a
 CC leucine to proline amino acid substitution at residue position 33.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-Infarct dementia.

SO Sequence 788 AA:

Query Match 99.0%; Score 4112; DB 21; Length 788;
 Best Local Similarity 99.2%; Pred. No. 2e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCGOCLAVSPMCAMCSDALPLGSPRODLKENLKDNCAPESIEFPYSE 60
 DB 27 GPNICTRGVSSCGOCLAVSPMCAMCSDALPLGSPRODLKENLKDNCAPESIEFPYSE 86
 QY 61 ARVLEDRPLSDKSGSDSQVTVSPQRIALRLRPDSDKNFSIQVQVEDYPVDIYLLMDL 120
 DB 87 ARVLEDRPLSDKSGSDSQVTVSPQRIALRLRPDSDKNFSIQVQVEDYPVDIYLLMDL 146
 QY 121 SYSKMDLMSIQNLGKTLATQMRKLTLSNRIGFAGAVDQVSPYWIISPEALENPTCYM 180
 DB 147 SYSKMDLMSIQNLGKTLATQMRKLTLSNRIGFAGAVDQVSPYWIISPEALENPTCYM 206
 QY 181 KTKVCLPMFGYKHNLTLDQVTRFNEEVKRSKOSRRNRDAEGGDALMQATVDEKIGMKN 240
 DB 207 KTKVCLPMFGYKHNLTLDQVTRFNEEVKRSKOSRRNRDAEGGDALMQATVDEKIGMKN 266
 QY 241 DASHLVEFTTDAKTALDGLAGIVQPNDCQCHWSDHMSASTTMDYPSGLMTEKLS 300
 DB 267 DASHLVEFTTDAKTALDGLAGIVQPNDCQCHWSDHMSASTTMDYPSGLMTEKLS 326
 QY 301 QKINILFAVTENVNLXQYSELIRGTTVGVLSMDSSNVQLIYDAYKIRSKVLELYR 360
 DB 327 QKINILFAVTENVNLXQYSELIRGTTVGVLSMDSSNVQLIYDAYKIRSKVLELYR 386
 QY 361 DLPEELSLSPNATCLNNEVTPGLKSCMGLKIGDVSFSEAVKRGCPQKESFTTKPVG 420
 DB 387 DLPEELSLSPNATCLNNEVTPGLKSCMGLKIGDVSFSEAVKRGCPQKESFTTKPVG 446
 QY 421 FKDSLIVQVTFDDCAQQAQAEFNSHRCNNNGTFFECGVRCGPGWLGSCQCESEEDYRP 480
 DB 447 FKDSLIVQVTFDDCAQQAQAEFNSHRCNNNGTFFECGVRCGPGWLGSCQCESEEDYRP 506
 QY 481 SQODECSPREGQPVQSGRGECLGQCVCVCHSSDFGKITGTYCECDDFSCVRYRGEMCSHG 540
 DB 507 SQODECSPREGQPVQSGRGECLGQCVCVCHSSDFGKITGTYCECDDFSCVRYRGEMCSHG 566
 QY 541 QCSGCGICLDSMDVGYCNCNTRFTDTCMSSNGLSGRGRCEGSCVCIOPGSDGTCKR 600
 DB 567 QCSGCGICLDSMDVGYCNCNTRFTDTCMSSNGLSGRGRCEGSCVCIOPGSDGTCKR 626
 QY 601 CPICPACTFKEKCEVCKKFDKALHDENTCNRYCDEIESVLELKDGTGKDAVNCYKNE 660
 DB 627 CPICPACTFKEKCEVCKKFDKALHDENTCNRYCDEIESVLELKDGTGKDAVNCYKNE 686

QY 661 DDCVVRQYVEDSSGSKSLTVVEEPCPKGPDIIVLISWGAIIILGLAALLIKLIT 720
 DB 687 ddcvvrqyvedssgsksltvveeepckgpdilvliswgailllglaalliklilt 746
 QY 721 IHDRKEFAKFEERAKRAMPDANNPLYKEATSPFTNTTYRGT 762
 DB 747 Ihtkefakfeeararakwtdannplykeatsftntltyrgt 788

RESULT 11
 AAY92452
 ID AAY92452 standard; Protein; 762 AA.

XX AAY92452;
 AC 10-AUG-2000 (first entry)
 DT 10-AUG-2000 (first entry)
 XX GPIIa variant Arg214Trp.
 KW GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KW allantoicins; platelet antigen; P1A1; P1A2; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.
 XX WO200020634-A1.
 PN 13-APR-2000.
 PD 01-OCT-1999; 99WO-1B01696.
 PF 01-OCT-1998; 98US-0102624.
 PR (NOVA-) NOVA MOLECULAR INC.
 PA Schappert K;
 PI WPI: 2000-303801/26.

XX Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIa or GPIIb locus and
 PT determining presence of variant GPIIa or GPIIb allele or isoform
 PS Disclosure; Page -: 55pp; English.

XX Known polymorphisms in GPIIa that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GPIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-Infarct dementia.

CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIa sequence which appears in Figure 3.

SO Sequence 762 AA:

Query Match 99.0%; Score 4111; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2.3e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCGOCLAVSPMCAMCSDALPLGSPRODLKENLKDNCAPESIEFPYSE 60
 DB 1 GPNICTRGVSSCGOCLAVSPMCAMCSDALPLGSPRODLKENLKDNCAPESIEFPYSE 60
 QY 61 ARVLEDRPLSDKSGSDSQVTVSPQRIALRLRPDSDKNFSIQVQVEDYPVDIYLLMDL 120

```

Db 61 arVledrplskgsdsgsyqvtvspqialrlrpdsskntsiqrvvedypvdiylmld 120
Qy 121 SYSMKDIWSIQNTGKTLATQMRKLTSLNRIGGAFVDRPVSYMISPEALENCYDM 180
Db 121 SYSMKDIWSIQNTGKTLATQMRKLTSLNRIGGAFVDRPVSYMISPEALENCYDM 180
Qy 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPGCGDAMQATVCDEKIGMWRN 240
Db 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPGCGDAMQATVCDEKIGMWRN 240
Qy 241 DASHLVEFTTDAKTHIALDGLAGIVQPNDCQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Db 241 DASHLVEFTTDAKTHIALDGLAGIVQPNDCQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Qy 301 OKNINLIFAVTENVNLYQNTSELIPGTIVGLSMSSNVQLIIVDAVGRKRSKVELEVR 360
Db 301 OKNINLIFAVTENVNLYQNTSELIPGTIVGLSMSSNVQLIIVDAVGRKRSKVELEVR 360
Qy 361 DLPEELISFNATCLNNEVTPGLKSCMGKIGDTPVSFSLIAKVRGCPQEKESFTTKPVG 420
Db 361 DLPEELISFNATCLNNEVTPGLKSCMGKIGDTPVSFSLIAKVRGCPQEKESFTTKPVG 420
Qy 421 FKSLIVQVTFDDCACAOAEPNSHRNCNNGNCTFECGVCRCGPGWLGSCQCESEEDYRP 480
Db 421 FKSLIVQVTFDDCACAOAEPNSHRNCNNGNCTFECGVCRCGPGWLGSCQCESEEDYRP 480
Qy 481 SQDECSPREGQPVCSQRCGECGQCCHVSHSDPFGKITGKYCECDDDSQVRYRGEMCSHG 540
Db 481 SQDECSPREGQPVCSQRCGECGQCCHVSHSDPFGKITGKYCECDDDSQVRYRGEMCSHG 540
Qy 541 QCSGDCDLSDMTGYCNCCTRTTDFTCMSSNGLICSGRKCCEGSCVC10PGSYGDTCEK 600
Db 541 QCSGDCDLSDMTGYCNCCTRTTDFTCMSSNGLICSGRKCCEGSCVC10PGSYGDTCEK 600
Qy 601 CPTRPACATFKKCECKEVEKRDALHDEMTNRCYCRDELSVKELKDTOKDAVNCTYKNE 660
Db 601 CPTRPACATFKKCECKEVEKRDALHDEMTNRCYCRDELSVKELKDTOKDAVNCTYKNE 660
Qy 661 DDCVAFQYEDSSGKSILYVEEPCPGPDLIVVLVSMGAILLIGLAALLIKMLTIT 720
Db 661 DDCVAFQYEDSSGKSILYVEEPCPGPDLIVVLVSMGAILLIGLAALLIKMLTIT 720
Qy 721 IHDRKEPAKFEERARAKWDTANNPYKCATSTFTNITYRG 762
Db 721 IHDRKEPAKFEERARAKWDTANNPYKCATSTFTNITYRG 762

```

RESULT 12

ID AAY92454 standard; Protein; 762 AA.

AC AAY92454;

DT 10-AUG-2000 (first entry)

DE GPIIIa variant Pro407Ala.

XX GPIIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;

XX allantigen; platelet antigen; PLAI; PLI2; Alzheimer's disease; stroke;

XX neurofibromatosis; Huntington's disease; depression; Parkinson's disease;

XX amyotrophic lateral sclerosis; multiple sclerosis; dementia.

XX Homo sapiens.

XX WO200020634-A1.

XX 13-APR-2000.

XX 01-OCT-1999; 99WO-IB01696.

XX 01-OCT-1998; 98US-0102624.

```

PA (NOVA-) NOVA MOLECULAR INC.
XX Schappert K:
XX WPI: 2000-303801/26.
XX
XX Identifying a subject at risk for a neurological disease comprises
XX determination of genotype or phenotype of GPIIIa or GPIIb locus and
XX determining presence of variant GPIIIa or GPIIb allele or isoform
XX
XX Disclosure: Page -: 55pp; English.
XX
XX Known polymorphisms in GPIIIa that may be determined to be variants
XX using the methods of the invention include AAY92447-56, which are variant
XX GPIIIa sequences derived from the mature protein.
XX Identifying a subject at risk for a neurological disease comprises
XX determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
XX subject, and determining the presence of a variant GPIIIa or GPIIb
XX allele or isoform, where the presence indicates an increased risk of
XX neurological disease. The method is useful for identifying a risk,
XX diagnosing and treating a neurological disease comprising Alzheimer's
XX disease, neurofibromatosis, Huntington's disease, depression,
XX amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
XX disease and multi-infarct dementia.
XX Note: This sequence is not given in the specification it was created
XX from the wild type human GPIIIa sequence which appears in Figure 3.
XX
XX Sequence 762 AA:

```

Query Match 99.0%; Score 4111; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2,3e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Qy 1 GPNCTTTRGVSSCCQCLAVSPMCAMCSDEALPISPRCDLKENLKNCAPESTEFYSE 60
Db 1 gpnctttrgvssccqclavspmcawcsdealpisprrcdlkenlkncapestiefyse 60
Qy 61 ARVLEDRLPSKSGSDSSQVQVSPORIALRLRPDSSKNFSIOVRQVEDYPVDIYYLMDL 120
Db 61 arVledrplskgsdsgsyqvtvspqialrlrpdsskntsiqrvvedypvdiylmld 120
Qy 121 SYSMKDIWSIQNTGKTLATQMRKLTSLNRIGGAFVDRPVSYMISPEALENCYDM 180
Db 121 SYSMKDIWSIQNTGKTLATQMRKLTSLNRIGGAFVDRPVSYMISPEALENCYDM 180
Qy 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPGCGDAMQATVCDEKIGMWRN 240
Db 181 KTTCLPMFGYKHYLTLDQVTRFENEYKKSQSVNRDAPGCGDAMQATVCDEKIGMWRN 240
Qy 241 DASHLVEFTTDAKTHIALDGLAGIVQPNDCQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Db 241 DASHLVEFTTDAKTHIALDGLAGIVQPNDCQCHVSDNHYASATTMDYPSLGLMTEKLS 300
Qy 301 OKNINLIFAVTENVNLYQNTSELIPGTIVGLSMSSNVQLIIVDAVGRKRSKVELEVR 360
Db 301 OKNINLIFAVTENVNLYQNTSELIPGTIVGLSMSSNVQLIIVDAVGRKRSKVELEVR 360
Qy 361 DLPEELISFNATCLNNEVTPGLKSCMGKIGDTPVSFSLIAKVRGCPQEKESFTTKPVG 420
Db 361 DLPEELISFNATCLNNEVTPGLKSCMGKIGDTPVSFSLIAKVRGCPQEKESFTTKPVG 420
Qy 421 FKSLIVQVTFDDCACAOAEPNSHRNCNNGNCTFECGVCRCGPGWLGSCQCESEEDYRP 480
Db 421 FKSLIVQVTFDDCACAOAEPNSHRNCNNGNCTFECGVCRCGPGWLGSCQCESEEDYRP 480
Qy 481 SQDECSPREGQPVCSQRCGECGQCCHVSHSDPFGKITGKYCECDDDSQVRYRGEMCSHG 540
Db 481 SQDECSPREGQPVCSQRCGECGQCCHVSHSDPFGKITGKYCECDDDSQVRYRGEMCSHG 540
Qy 541 QCSGDCDLSDMTGYCNCCTRTTDFTCMSSNGLICSGRKCCEGSCVC10PGSYGDTCEK 600
Db 541 QCSGDCDLSDMTGYCNCCTRTTDFTCMSSNGLICSGRKCCEGSCVC10PGSYGDTCEK 600

```

QY 601 CPTCPDCTFKECKEYCKEDRGALHDETCNRCRDEIESVKELKDTGKDAVNCTYKNE 660
 DB 601 cptcpdactfkeckevckkfdrpymtentcnrycrdeiesvkelkdtgkdavncitykne 660
 QY 661 DDCVVRFGYEDSSGSKSLIYVVEEPECPKGPDLIVLLSVGAILLIGLALLIWKLLIT 720
 DB 661 ddcvvrfgyedsqsksllyvveepecpkgpdlivllsvmgaillliglaalllwkllit 720
 QY 721 IHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNTITYRGT 762
 DB 721 ihdrekefakfeeararawkdtannplykeatsftntityrgt 762

RESULT 13

AAV92455
 ID AAV92455 standard; Protein: 762 AA.

AAV92455:

10-AUG-2000 (first entry)

GPIIa variant Arg636Cys.

KW GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KX allonantigen; platelet antigen; PLA1; PLA2; Alzheimers disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

PE 01-OCT-1999: 99WO-1B01696.

PR 01-OCT-1998: 98US-0102624.

PA (NOVA-) NOVA MOLECULAR INC.

PI Schappert K;

DR WPI: 2000-303801/26.

PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIa or GPIIb locus and
 PT determining presence of variant GPIIa or GPIIb allele or isoform

PS Disclosure: Page -: 55pp; English.

CC known polymorphisms in GPIIa that may be determined to be variants
 CC using the methods of the invention include AAV92447-56, which are variant
 CC GPIIa sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntingdon's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIa sequence which appears in Figure 3.

XX Sequence 762 AA;

Query Match 99.0%; Score 4111; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2,3e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSQOCCLAASPNCACSDALPLGSPKCDLKNLNDNCAPESIEEPVSE 60
 DB 1 gpnicttrgvssccqclavspmcawcsdeallplgspcdclkenllkdnccapesiefpvse 60
 QY 61 ARVLEDPRLSDKSGSDSSQVTVSPQRIALRLRPDDSKNFSIGVROVEDYPVDIYVLMDL 120
 DB 61 arvledrplsdksqsdssqvtvqspqriralrlrppddsknfsigvrvvedypvdiylmld 120
 QY 121 SYSMKDDLMSIQNLGTYLATQMRKLSNLRIGFCAFYDKPVSPEYKISPEALENPECYDM 180
 DB 121 sysmkddlmsiqnlgtylatqmrklsnlrigfcaydkpvspeykispealenpocydm 180
 QY 181 KTTCLPMFGRKHVLTLDQVTRFNEEYKOSYSNNRPAPEGCDATVQATVCEKIGMWRN 240
 DB 181 ktclpmlfgykhvltldqvtrfneevkqsvsnrdapegidaImqatvcdexlqwrn 240
 QY 241 DASHLVFTTDAKTHIALDGRLAGIYVQPNQOCHVSGSDNHSYSTWDYPSLGLMEFKLS 300
 DB 241 dashllvftcdakthialdgrlagivqpnqochvsgsdnhsystndypslglmckls 300
 QY 301 QKNINLIFAVTENVNLYXNSELIPGTVGVLSMDSSNVLQILVDAYGKIRSKVELEVR 360
 DB 301 qkninlifavtenvnlyqnsellpgltvgvlsmssnvlqilvdaygkirkvelevr 360
 QY 361 DLPEELISFPAATLNNENVPGLKSCMGKRIQDVFVSIFAKVRGCRQEKESTTIRPYG 420
 DB 361 dlpeelistsfpaatlneevlpglkscmglikgdtvvsifaeakvrgcrqekestsflkpyg 420
 QY 421 FKDSLIVQVTFPDCCACAOAEPNSHRCNNGNFGFEGVCVCGPGMGLSCQCECSEEDYRP 480
 DB 421 fkdslivqvtfdccacacqaqepnshrcnngnfgfcgvrcrpgmglsgqceceedyrp 480
 QY 481 SQODECSPREGQPYVCSORGECLCGQVCYCHSSDFGKITGKYCECDPFCVRYKGEKSGHG 540
 DB 481 sqgdecspregqpyvcsgqgeclcgqvcvchssdfgkitgkycecdpfcvrykgemcshg 540
 QY 541 QCSGDCCLCDSDDMTGYCNCITRTDTCMSSNGLLCSGRGKCECCSCYCIOPGSGDICEK 600
 DB 541 qcsgdcclcdsddmtgycncitrtctcmssngllcsgrgkceccscvcilp9sygdicek 600
 QY 601 CPTCPDCTFKECKEYCKEDRGALHDETCNRCRDEIESVKELKDTGKDAVNCTYKNE 660
 DB 601 cptcpdactfkeckevckkfdrpymtentcnrycrdeiesvkelkdtgkdavncitykne 660
 QY 661 DDCVVRFGYEDSSGSKSLIYVVEEPECPKGPDLIVLLSVGAILLIGLALLIWKLLIT 720
 DB 661 ddcvvrfgyedsqsksllyvveepecpkgpdlivllsvmgaillliglaalllwkllit 720
 QY 721 IHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNTITYRGT 762
 DB 721 ihdrekefakfeeararawkdtannplykeatsftntityrgt 762

RESULT 14

AAV92449
 ID AAV92449 standard; Protein: 762 AA.

AAV92449:

10-AUG-2000 (first entry)

GPIIa variant Asp118Tyr.

KW GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KX allonantigen; platelet antigen; PLA1; PLA2; Alzheimers disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.

OS Homo sapiens.

PN WO200020634-A1.

PD 13-APR-2000.

XX 01-OCT-1999: 99WO-IB01696.
 XX
 XX 01-OCT-1998: 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K:
 XX WPI: 2000-303801/26.
 DR
 XX
 PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIIa or GPIIb locus and
 PT determining presence of variant GPIIIa or GPIIb allele or isoform
 XX
 PS Disclosure; Page -: 55pp; English.
 XX
 CC Known polymorphisms in GPIIIa that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GPIIIa sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIIa sequence which appears in Figure 3.
 CC
 XX
 SQ Sequence 762 AA:

Query Match 98.9%; Score 4110; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2.7e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDALPLGSPRCDLKENLKDNCAPESIEFPVSE 60
 DB 1 gpnictrgvsscgqlavspmcawcsdealplgsprcdlkenllkdncaesiefpvse 60
 QY 61 ARYLERPISDKGSDSSQVTOVSFORIALRLPDDSKNFSIQVROVEDYPVDIYLMDL 120
 DB 61 aryledrplisdqsgsdsvltqvsfpjalrlrpdsknfsiqvgyvedypvdiylmlyl 120
 QY 121 SYSMKDDLSIONLGTKLTOMRKLTSLNRIGFAFVOKPSPVYVISPPEALENPPCYDM 180
 DB 121 sysmkddlwsigqltklctgmkltsnlrfgatvdkpvsppymysppealenpcydm 180
 QY 181 KTTCLPMFGYKHVLTLDVTRFNEEVKKOSVSRNKDAPEGGFDAIMQATVDEKIGRMN 240
 DB 181 ktclpmfgykhvltldvtrfneevkkosvsnrndapeggsfdaimatvdekigymn 240
 QY 241 DASHLVFTTDAKTHIALDGLAGIVQPNQGCHVGSNDHNSASTTMDYPSLGLMTERKLS 300
 DB 241 dashlvfttdakthialdglagivqpdgqchvgsndhysasttmdypslglmtkls 300
 QY 301 GKNINLIPAVTENYVNLVQYSELIPGTTVGVLSMDSNVQLIYDAIGKIRSKYLEVR 360
 DB 301 gkninlilavtenvnlvqnyseilpgttvgvlsmdssnvqlivdaygkirkyleavr 360
 QY 361 DIPPEELISFNATCLNNEVTPGLKSCMGKIDTVSFSIEAKVRCCPEKEKSFIRKPVG 420
 DB 361 dippeelistsfnatclnnevtplgkscmgkldtvsvfsieakvrycpekeksfirkpv 420
 QY 421 FFDLSLIVQTFDCCACQAQAPNSHRNNGNGTRECVCVCRGPGWLSQCECEDYR 480
 DB 421 ffdslivqtfddccacqaqapnshrcnngngtrecvcvcrpgwlsqcecedyrp 480
 QY 481 SQDDCSPREGQPVCSQGECLCGCVCHSSDFGKITRKYCECDDFSCVRRKGEKMSGHG 540

DB 481 sqddcspreggpvcsqrgeclcgqcvchssdfgkitkycecddfsuvrykmgemcsgh 540
 QY 541 QCSGDCICLDSWMTGYTCNCTTRDTGMSNGLCSGKKECCSCVCTPGSTGDTCEK 600
 DB 541 qcsgcdicldswwtgytcncttrdtgmsnlgcsqkccscvcvctpgsgtdce 600
 QY 601 CPTCPDCTFFKCEVCEKFKKFRGALHDENTNRCRCRDIIESEVKELEKDPKAVACTYKNE 660
 DB 601 cptcpdctffkceveckfkfdrepmtenctnrycrdesvkelkdygkavactykne 660
 QY 661 DDCVVRFOYEDSSGKSLIYVEEPCPKGPDILVLLSVGAILLGLAALLIWLKLLIT 720
 DB 661 ddcvvrfoyedssgksliyyveeepckgpdilvllsvmgailllglaalliwllit 720
 QY 721 IHDRKEFAKFEERARAKWDTANPNLYKEATSTFNTRYRGT 762
 DB 721 ihdrkefakfeerarakwdtanpnlykeatstfntrygt 762

RESULT 15
 ID AAY92453 standard; Protein: 762 AA.
 AAY92453
 AC AAY92453;
 XX 10-APR-2000 (first entry)
 DT
 XX
 DE GPIIIa variant Cys374Tyr.
 KW GPIIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
 KW allolantigen; platelet antigen; PLAI; PLAZ; Alzheimer's disease; stroke;
 KW neurofibromatosis; Huntingdon's disease; depression; Parkinson's disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX
 OS Homo sapiens.
 XX
 PN WO200020634-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999: 99WO-IB01696.
 XX
 PR 01-OCT-1998: 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K:
 XX WPI: 2000-303801/26.
 DR
 XX
 PT Identifying a subject at risk for a neurological disease comprises
 PT determination of genotype or phenotype of GPIIIa or GPIIb locus and
 PT determining presence of variant GPIIIa or GPIIb allele or isoform
 XX
 PS Disclosure; Page -: 55pp; English.
 XX
 CC Known polymorphisms in GPIIIa that may be determined to be variants
 CC using the methods of the invention include AAY92447-56, which are variant
 CC GPIIIa sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIIa or GPIIb locus of a
 CC subject, and determining the presence of a variant GPIIIa or GPIIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, Huntington's disease, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 CC from the wild type human GPIIIa sequence which appears in Figure 3.
 CC
 XX
 SQ Sequence 762 AA:

Query Match 98.9%; Score 4108; DB 21; Length 762;
Best Local Similarity 99.2%; Pred. No. 3.7e-282;
Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGVSSCQOCLAVSPKACNSDEALPLGSPRCOLKENLKDNCAPESIEFPVSE 60
DB 1 gpnicttrgvsscqqclavspmcawcsdealplgsprcdlkenllkdncapestiefpvse 60
QY 61 ARVLEDRPLSDKSGSSQVTOVSQRIALRLRPDSSKNFSIQRVQVEDYPVDIYYLMDL 120
DB 61 arvledrplsdksqsgdsqvtqvsqprlrlrtpdssknfsiqrvqedypvdiyylmql 120
QY 121 SYAMKDLWMSIQNLGTRKLTQMRKLTSLNLRIGFAFVDKPVSPMYISPEALENPCYDM 180
DB 121 syamkddlwsiqnlgtkrlatqmrkltslnlrigfafvdkpvspmyispealenpcydm 180
QY 181 KTTCLPMEGKHYVLTLDQYTRFNEVEYKKOSVSRNRDAPEGGFDAIMQATVCDEKIGWRN 240
DB 181 ktclpmeqkhyvltldqytrfneevkksvsnrdapeggfdaimqatvcdekigwrn 240
QY 241 DASHLVFTTDAKTHIALDRLAGIVOPNDGCHVSDNHYSASTMDYPSLGLMTEKLS 300
DB 241 dashlvfttdakthialdgrlvgpnbgqchvsdnhysastcmdypslglmtekls 300
QY 301 QKINILFAVTENVNLYQYNSLIPGTGVLSDMSNVYLQILVDAYGKIRKVELEVYR 360
DB 301 qkinilfavtenvnlqyyselipgtvgvlsmdssnvylqilvdaygkirkvelevyr 360
QY 361 DLPEELSLSFNATCLNNNEVIFGLKSCMGLKIGPTVSFSTIAKVRGCPQEKESFTIKPVC 420
DB 361 dlpeelslsfnatclnnnevifglkscmglkigptvsfstiakvrgcpqekesftikpvc 420
QY 421 FKDSLIVQVTFPDCACQAOAEPNSHRNNGNGTFEGGVCRCGPGWLGSGCESEEDYRP 480
DB 421 fkdslivqvtfpdcacqaepnshrcnngngtfeggvcrcgpgwlgsgceceedyrp 480
QY 481 SQQDECSPRREGQPYCSGRGECGCGVCHSSDFGKITGKTCBCDDFSCVRYKGBMCSGHG 540
DB 481 sqqdecsprregqpycsgrgeclcgqcvchssdfgkltgkyccecdtscvrykgemcsghg 540
QY 541 QCSGDCGLDSDMTGYCNCCTTRDTMCSSNGLLCSGRKCEGSCVCITQPGSYGDPCEK 600
DB 541 qcsgdcglcdsdmtgycnccttrdtmcssngllcsgrgkcegsccvcitqpsygdpcck 600
QY 601 CPTCPDACEFKKECECKKFDRCALHDENTCNRYCRDEISYVELKDGTGKDAVNCITYKNE 660
DB 601 cptcpdactfkkececkkfdrepmtenconycrdeiesvkelkdtgkdavncitykne 660
QY 661 DDCVYRQYTEDSSGKSILYVEEPECCKGPDILYVLLSYGAILLGLAALLIKLLIT 720
DB 661 ddcvyrqytedssgksilyveepecckgpdilvllsygailllglaallikllit 720
QY 721 IHDKREFAKFEERARAKWTDANNPLYKEATSTFTNTITYRGT 762
DB 721 ihdkrefakfeerarakwdannplykeatstftntityrgt 762

Search completed: May 19, 2002, 10:38:56
Job time: 9118 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:07:51 ; Search time 36.84 Seconds
(without alignments)
505.220 Million cell updates/sec

Title: US-09-673-302a-1
Sequence: 1 GPNICTRGVSSCQCCLAVS.....NNPLYKATSTFTNITYRGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5a_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6a_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6b_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6c_COMB.pep:*
7: /cgn2_6/ptodata/2/1aa/6d_COMB.pep:*
8: /cgn2_6/ptodata/2/1aa/6e_COMB.pep:*
9: /cgn2_6/ptodata/2/1aa/6f_COMB.pep:*
10: /cgn2_6/ptodata/2/1aa/6g_COMB.pep:*
11: /cgn2_6/ptodata/2/1aa/6h_COMB.pep:*
12: /cgn2_6/ptodata/2/1aa/6i_COMB.pep:*
13: /cgn2_6/ptodata/2/1aa/6j_COMB.pep:*
14: /cgn2_6/ptodata/2/1aa/6k_COMB.pep:*
15: /cgn2_6/ptodata/2/1aa/6l_COMB.pep:*
16: /cgn2_6/ptodata/2/1aa/6m_COMB.pep:*
17: /cgn2_6/ptodata/2/1aa/6n_COMB.pep:*
18: /cgn2_6/ptodata/2/1aa/6o_COMB.pep:*
19: /cgn2_6/ptodata/2/1aa/6p_COMB.pep:*
20: /cgn2_6/ptodata/2/1aa/6q_COMB.pep:*
21: /cgn2_6/ptodata/2/1aa/6r_COMB.pep:*
22: /cgn2_6/ptodata/2/1aa/6s_COMB.pep:*
23: /cgn2_6/ptodata/2/1aa/6t_COMB.pep:*
24: /cgn2_6/ptodata/2/1aa/6u_COMB.pep:*
25: /cgn2_6/ptodata/2/1aa/6v_COMB.pep:*
26: /cgn2_6/ptodata/2/1aa/6w_COMB.pep:*
27: /cgn2_6/ptodata/2/1aa/6x_COMB.pep:*
28: /cgn2_6/ptodata/2/1aa/6y_COMB.pep:*
29: /cgn2_6/ptodata/2/1aa/6z_COMB.pep:*
30: /cgn2_6/ptodata/2/1aa/6aa_COMB.pep:*
31: /cgn2_6/ptodata/2/1aa/6ab_COMB.pep:*
32: /cgn2_6/ptodata/2/1aa/6ac_COMB.pep:*
33: /cgn2_6/ptodata/2/1aa/6ad_COMB.pep:*
34: /cgn2_6/ptodata/2/1aa/6ae_COMB.pep:*
35: /cgn2_6/ptodata/2/1aa/6af_COMB.pep:*
36: /cgn2_6/ptodata/2/1aa/6ag_COMB.pep:*
37: /cgn2_6/ptodata/2/1aa/6ah_COMB.pep:*
38: /cgn2_6/ptodata/2/1aa/6ai_COMB.pep:*
39: /cgn2_6/ptodata/2/1aa/6aj_COMB.pep:*
40: /cgn2_6/ptodata/2/1aa/6ak_COMB.pep:*
41: /cgn2_6/ptodata/2/1aa/6al_COMB.pep:*
42: /cgn2_6/ptodata/2/1aa/6am_COMB.pep:*
43: /cgn2_6/ptodata/2/1aa/6an_COMB.pep:*
44: /cgn2_6/ptodata/2/1aa/6ao_COMB.pep:*
45: /cgn2_6/ptodata/2/1aa/6ap_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4119	99.2	788	2	US-07-728-215-32
2	4119	99.2	788	4	US-08-938-085A-32
3	3768	90.7	718	1	US-08-444-792-4
4	3768	90.7	718	1	US-08-445-042-4
5	2339	56.3	799	1	US-08-054-077C-2
6	2156	51.9	788	2	US-07-728-215-27
7	2156	51.9	788	4	US-08-938-085A-27
8	1831	44.1	798	2	US-07-728-215-30
9	1831	44.1	798	4	US-08-938-085A-30
10	1676	40.3	577	2	US-07-728-215-29
11	1676	40.3	577	4	US-08-938-085A-29
12	1535.5	37.0	846	4	US-07-728-215-33
13	1535.5	37.0	846	4	US-08-938-085A-33
14	1535.5	36.4	769	2	US-08-789-078-1
15	1535.5	36.4	769	2	US-08-752-633-1
16	1535.5	36.4	769	2	US-08-476-062A-45
17	1535.5	36.4	769	2	US-07-728-215-31
18	1535.5	36.4	769	4	US-08-938-085A-31
19	1535.5	36.4	769	5	PCT-US95-04886-1
20	1535.5	36.4	769	5	PCT-US96-01314-45
21	1343.5	32.3	676	3	US-08-630-172-10
22	1343.5	32.3	676	4	US-09-375-419-10
23	1335	32.0	255	5	PCT-US96-01314-57
24	1335	32.0	255	5	US-08-454-455-4
25	1326.5	31.9	769	1	US-08-454-455-6
26	1278	30.8	266	2	US-07-808-457-3
27	1278	30.8	266	2	US-07-808-457-3

28	1278	30.8	266	5	PCT-US92-10178-3	Sequence 3, App1
29	1271	30.6	266	2	US-07-808-457-5	Sequence 5, App1
30	1271	30.6	266	5	PCT-US92-10178-5	Sequence 5, App1
31	1065	25.6	200	1	US-08-170-596-15	Sequence 15, App1
32	883.5	21.3	256	5	PCT-US96-01314-55	Sequence 55, App1
33	797	19.2	235	5	PCT-US96-01314-54	Sequence 54, App1
34	731	17.8	344	4	US-09-393-554-1	Sequence 1, App1
35	731	17.6	164	2	US-07-808-457-19	Sequence 19, App1
36	731	17.6	164	4	PCT-US92-10178-19	Sequence 19, App1
37	730	17.6	344	4	US-09-393-554-4	Sequence 4, App1
38	728	17.5	344	4	US-09-393-554-3	Sequence 3, App1
39	726	17.5	344	4	US-09-393-554-5	Sequence 5, App1
40	724	17.4	164	2	US-07-808-457-29	Sequence 29, App1
41	724	17.4	164	4	PCT-US92-10178-29	Sequence 29, App1
42	722	17.4	344	4	US-09-393-554-2	Sequence 2, App1
43	688	16.6	251	5	PCT-US96-01314-59	Sequence 59, App1
44	656	15.8	150	2	US-07-808-457-17	Sequence 17, App1
45	656	15.8	150	5	PCT-US92-10178-17	Sequence 17, App1

ALIGNMENTS

RESULT 1
US-07-728-215-32
Sequence 32, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Ylto
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: AMINO ACID
STRANDEDNESS: Single
TOPOLOGY: linear
US-07-728-215-32

Query Match 99.2% Score 4119: DB 2: Length 788:
Best Local Similarity 99.3% Pred. No. 0:
Matches 757: Conservative 0: Mismatches 5: Indels 0: Gaps 0:
QY 1 GPNICTRGVSSCQCCLAVSPKCMKSDALPLGSPRCIDKENVLNDNCAPESIEPVS 60
DB 27 GPNICTRGVSSCQCCLAVSPKCMKSDALPLGSPRCIDKENVLNDNCAPESIEPVS 86

```

QY 61 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 120
D 87 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 146
QY 121 SYKMKDLMISIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 180
D 147 SYKMKDLMISIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 206
QY 181 KTTCLPMEGKXHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 240
D 207 KTTCLPMEGKXHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 266
QY 241 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 300
D 267 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 326
QY 301 OKNINLIFAVTENNVNLYONTSSELIPGTIVGLSMDSNNVLOLIVDAYGKIRSKVELEVR 360
D 327 OKNINLIFAVTENNVNLYONTSSELIPGTIVGLSMDSNNVLOLIVDAYGKIRSKVELEVR 386
QY 361 DLPEELISFNATCLNNEVNLGKSCMGKIGDTVSPFISIAKVRGCPQEKESFTTKPVG 420
D 387 DLPEELISFNATCLNNEVNLGKSCMGKIGDTVSPFISIAKVRGCPQEKESFTTKPVG 446
QY 421 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNGNGTFEECGVCRGPGWLGSCCESEEDYRP 480
D 447 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNGNGTFEECGVCRGPGWLGSCCESEEDYRP 506
QY 481 SQODECSPREGQPVCSORGECLCGQCHVSDNHYSASTMDYPSLGLMTEKLS 540
D 507 SQODECSPREGQPVCSORGECLCGQCHVSDNHYSASTMDYPSLGLMTEKLS 566
QY 541 QCSGCDCLCSDMTGYCNCCTTRDTGMSNGLLCSGRKCEGSCVCIQPGSYGDTCEK 600
D 567 QCSGCDCLCSDMTGYCNCCTTRDTGMSNGLLCSGRKCEGSCVCIQPGSYGDTCEK 626
QY 601 CPTCPDACTFKKECEVECKKFRGALHDENTCNRYCRDELESYVELDGTGDAVNCTYKNE 660
D 627 CPTCPDACTFKKECEVECKKFRGALHDENTCNRYCRDELESYVELDGTGDAVNCTYKNE 686
QY 661 DDCVVRQYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLIGLAALLIMKLLIT 720
D 687 DDCVVRQYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLIGLAALLIMKLLIT 746
QY 721 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITYRGT 762
D 747 IHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNITYRGT 788

```

```

RESULT 2
US-08-938-085A-32
; Sequence 32, Application US/08938085A
; Patent No. 6339148
; GENERAL INFORMATION:
; APPLICANT: Shepard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytela, Robert
; TITLE OF INVENTION: A No. 6339148e1 Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938, 085A
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728, 215
; FILING DATE: 11-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parent, Annette S
; REGISTRATION NUMBER: 42,058
; REFERENCE/DOCKET NUMBER: 023070-080210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-938-085A-32

```

```

Query Match 99.2%; Score 4119; DB 4; Length 788;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GNRICTRQVSSCQOCLAVSPKAMCSDBALPLSGPRCDLKENLLKNCAPSEIEFPVSE 60
D 27 GNRICTRQVSSCQOCLAVSPKAMCSDBALPLSGPRCDLKENLLKNCAPSEIEFPVSE 86
QY 61 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 120
D 87 ARVLEDRPLSDKSGSDSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYVPVDIYYIMDL 146
QY 121 SYKMKDLMISIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 180
D 147 SYKMKDLMISIQNLGTLATQMRKLTSLNLRIGFAGVDPKPYSPYMTISPEALENPCYDM 206
QY 181 KTTCLPMEGKXHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 240
D 207 KTTCLPMEGKXHVLTLDQVTRFNEEYKOSVSRNRDAPRGFPALIMQATVCEKIGWRN 266
QY 241 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 300
D 267 DASHLVEFTTDAKTHIALDGRLAGIYOPNDGQCHVSDNHYSASTMDYPSLGLMTEKLS 326
QY 301 OKNINLIFAVTENNVNLYONTSSELIPGTIVGLSMDSNNVLOLIVDAYGKIRSKVELEVR 360
D 327 OKNINLIFAVTENNVNLYONTSSELIPGTIVGLSMDSNNVLOLIVDAYGKIRSKVELEVR 386
QY 361 DLPEELISFNATCLNNEVNLGKSCMGKIGDTVSPFISIAKVRGCPQEKESFTTKPVG 420
D 387 DLPEELISFNATCLNNEVNLGKSCMGKIGDTVSPFISIAKVRGCPQEKESFTTKPVG 446
QY 421 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNGNGTFEECGVCRGPGWLGSCCESEEDYRP 480
D 447 FKDSLIVQVTFEDCDCAQOAEPNSHRCNNGNGTFEECGVCRGPGWLGSCCESEEDYRP 506
QY 481 SQODECSPREGQPVCSORGECLCGQCHVSDNHYSASTMDYPSLGLMTEKLS 540
D 507 SQODECSPREGQPVCSORGECLCGQCHVSDNHYSASTMDYPSLGLMTEKLS 566
QY 541 QCSGCDCLCSDMTGYCNCCTTRDTGMSNGLLCSGRKCEGSCVCIQPGSYGDTCEK 600
D 567 QCSGCDCLCSDMTGYCNCCTTRDTGMSNGLLCSGRKCEGSCVCIQPGSYGDTCEK 626
QY 601 CPTCPDACTFKKECEVECKKFRGALHDENTCNRYCRDELESYVELDGTGDAVNCTYKNE 660
D 627 CPTCPDACTFKKECEVECKKFRGALHDENTCNRYCRDELESYVELDGTGDAVNCTYKNE 686
QY 661 DDCVVRQYEDSSGKSLIYVEEPECKPDPDILVLLSWGAILLIGLAALLIMKLLIT 720

```


DB 687 DDCVVRFOYEDSSGKSLIYVEEPCPKGPDLVLLVSMGAILLIGLAALLIMKLLIT 746
QY 721 IHDRKEFAKEEERARAKMDTANNPLYKEATSTFTNITTYRGT 762
DB 747 IHDRKEFAKEEERARAKMDTANNPLYKEATSTFTNITTYRGT 788

RESULT 3

US-08-444-792-4
Sequence 4, Application US/08444792
Patent No. 5726037
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Napier, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
NUMBER OF SEQUENCES: 18
TITLE OF INVENTION: POLYPEPTIDES
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,792
FILING DATE: 19-May-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218878
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: P0552PLC304
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-444-792-4

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDPALPGSPRCDLKENLKNCAPESTIEPPVSE 60
Query Match 90.7%; Score 3768; DB 1; Length 718;
Best Local Similarity 99.3%; Pred. No. 2,5e-295;
Matches 687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 27 GPNICTRGVSSCOQCLAVSPMCAMCSDPALPGSPRCDLKENLKNCAPESTIEPPVSE 86
QY 61 ARVLEDRPLSDKSGSDSSQVTVSPQRIALRLRPDSKNFSIOVROVEDYVVDIYVMDL 120
DB 87 ARVLEDRPLSDKSGSDSSQVTVSPQRIALRLRPDSKNFSIOVROVEDYVVDIYVMDL 146
QY 121 SYSKMDLWSTONLGTKATQMRKLTSLNRIGFGAFVDPVSPYIYISPPALENPCYDM 180
DB 147 SYSKMDLWSTONLGTKATQMRKLTSLNRIGFGAFVDPVSPYIYISPPALENPCYDM 206
QY 181 KTTCLPMFGYKRVLTLPQVTRFNEEVKQOSVRRRDAPEGGFPAIMQATYCDKIGRN 240
DB 207 KTTCLPMFGYKRVLTLPQVTRFNEEVKQOSVRRRDAPEGGFPAIMQATYCDKIGRN 266
QY 241 DASHLVEFTDAKTHIALDGLAGIVQPNDCQCHVSDNHYASATYMDYPSLGLMTERLS 300
DB 267 DASHLVEFTDAKTHIALDGLAGIVQPNDCQCHVSDNHYASATYMDYPSLGLMTERLS 326
QY 301 QKNILIFAVTEVNVNLYQNSSELIPGTVGLSDSSNVQLYDVGKIRSKVELEVR 360
DB 327 QKNILIFAVTEVNVNLYQNSSELIPGTVGLSDSSNVQLYDVGKIRSKVELEVR 386
QY 361 DPEELISFNATCLNNEYIPGLKSCMGLKIGDYVFSIEAKVRGCPQEKESFTIKRPG 420
DB 387 DPEELISFNATCLNNEYIPGLKSCMGLKIGDYVFSIEAKVRGCPQEKESFTIKRPG 446
QY 421 FKDSLIVQVTFEDCCACQAQAEPNSHRCNNGNGTEFCGVCRCGPGWLCSCCESEEDYR 480
DB 447 FKDSLIVQVTFEDCCACQAQAEPNSHRCNNGNGTEFCGVCRCGPGWLCSCCESEEDYR 506
QY 481 SODDCSPREGQPVCSQRECLCGQCVCHSSDFGKITKCYCEDFSCVVRKKGEMCSHG 540
DB 507 SODDCSPREGQPVCSQRECLCGQCVCHSSDFGKITKCYCEDFSCVVRKKGEMCSHG 566
QY 541 QCSGCDLSDSDMTGYNCNCTRTDTCMSNGLCSGRGKEGSCGVC10P6SYDTEK 600
DB 567 QCSGCDLSDSDMTGYNCNCTRTDTCMSNGLCSGRGKEGSCGVC10P6SYDTEK 626
QY 601 CPTCPDCTFEKCEYCKKEDRGALHDENTCNRYCRDEIESYKELDKGDAVNTYRNE 660
DB 627 CPTCPDCTFEKCEYCKKEDRGALHDENTCNRYCRDEIESYKELDKGDAVNTYRNE 686
QY 661 DDCVVRFOYEDSSGKSLIYVEEPCPKGP 692
DB 687 DDCVVRFOYEDSSGKSLIYVEEPCPKGP 718

RESULT 4
US-08-445-042-4
Sequence 4, Application US/08445042
Patent No. 5726290
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Napier, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,042

```

FILING DATE: 19-May-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218878
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: P05522PIC3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-445-042-4

```

```

Query Match          90.7%: Score 3768; DB 1; Length 718;
Best Local Similarity 99.3%; Pred. No. 2.5e-295;
Matches 667; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSPRCDLKENLKLKNCAPESIEEPVSE 60
DB 27 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSPRCDLKENLKLKNCAPESIEEPVSE 86
QY 61 ARVLEDRPLSDKSGSSQVTVQVSPQRIALRLAPDDSKNFSIOVROYEDYVDIYIYLM 120
DB 87 ARVLEDRPLSDKSGSSQVTVQVSPQRIALRLAPDDSKNFSIOVROYEDYVDIYIYLM 146
QY 121 SYSKMDLWSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENCYDM 180
DB 147 SYSKMDLWSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENCYDM 206
QY 181 KTTCLPMEGKHYVLTLDQVTRNEEVKQSVSRNDABEGGFDAIMQATVCEDEKIGWN 240
DB 207 KTTCLPMEGKHYVLTLDQVTRNEEVKQSVSRNDABEGGFDAIMQATVCEDEKIGWN 266
QY 241 DASHLVETPDATHTALDGRLAGIYQPNDDGCHVSDHHSASTMDYPSGLMTEKLS 300
DB 267 DASHLVETPDATHTALDGRLAGIYQPNDDGCHVSDHHSASTMDYPSGLMTEKLS 326
QY 301 QKNINILFAVTENVNMLYQNSSELLPGTTGVLSMDSNVLIQIYDAVKISKVELEYR 360
DB 327 QKNINILFAVTENVNMLYQNSSELLPGTTGVLSMDSNVLIQIYDAVKISKVELEYR 386
QY 361 DLPEELSLSFNATCLNNEVYIPGLKSGMGLKIGDVSFSIEAKVRGCPQEKESFTIKPVG 420
DB 387 DLPEELSLSFNATCLNNEVYIPGLKSGMGLKIGDVSFSIEAKVRGCPQEKESFTIKPVG 446
QY 421 FDSLLVQVTFPDCCACQQAEPNSHRCKNNGTFFECGVCRCRGPGLSGCCSEEDYRP 480
DB 447 FDSLLVQVTFPDCCACQQAEPNSHRCKNNGTFFECGVCRCRGPGLSGCCSEEDYRP 506
QY 481 SQQDECSPREGQPVCSQRECECLGQGVCHSSDPFGKITGKCCDDPSCVRYKGEWCSGSH 540
DB 507 SQQDECSPREGQPVCSQRECECLGQGVCHSSDPFGKITGKCCDDPSCVRYKGEWCSGSH 566

```

```

QY 541 QCSGDCICDSDWMTGYCNCITRTDTOMSSNGLLCSGKGCSCGVCIIQPSYDPTCEK 600
DB 567 QCSGDCICDSDWMTGYCNCITRTDTOMSSNGLLCSGKGCSCGVCIIQPSYDPTCEK 626
QY 601 CPTCPDACTFKKECEVCKKFKDGRGALHDEWTCNRYCRDEIESYKELDKDGAIVNCTYKNE 660
DB 627 CPTCPDACTFKKECEVCKKFKDGRGALHDEWTCNRYCRDEIESYKELDKDGAIVNCTYKNE 686
QY 661 DDCVVRFOYEDSSGSKSLIYVEEPCPKGPD 692
DB 687 DDCVVRFOYEDSSGSKSLIYVEEPCPKGPD 718

```

RESULT 5

```

US-08-054-077C-2
Sequence 2, Application US/08054077C
Patent No. 5527679
GENERAL INFORMATION:
APPLICANT: HEMLER, MARTIN E.
APPLICANT: RAMASWAMY, HEMAVATHI
TITLE OF INVENTION: HUMAN INTEGRIN 5 SUBUNIT PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
ADDRESS: CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/054,077C
FILING DATE: 27-APR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/694314
FILING DATE: 01-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, DAVID S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 40937
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-054-077C-2

```

```

Query Match          56.3%: Score 2339; DB 1; Length 799;
Best Local Similarity 55.5%; Pred. No. 3.9e-180;
Matches 433; Conservative 118; Mismatches 205; Indels 24; Gaps 11;

```

```

QY 1 GPNICTRGVSSCOQCLAVSPMCAMCSDEALPLGSP-----RQDLKENLKLKNCAPESIE 55
DB 24 GINICTSGSANSCECECLIHKPCAMCSKE--DFGSPSRTSRCDLRLNLVKNCGGE-IE 80
QY 56 PVESEARVLEDRPLSDKSGSS--QVTVSPQRIALRLAPDDSKNFSIOVROYEDYVDI 114
DB 81 SPASSFHLRLPLSSKSGSSAGMDVIQMTQELAVNLRGDKTTFPOLQVROYEDYVDI 140
QY 115 YLMDLSYSKMDLWSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALE 174
DB 115 YLMDLSYSKMDLWSIQNLGTRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALE 174

```

```

Db 141 YLMDLSLMSKDDLDNIRSLGTLAEEMKRLTNSFRLGFGSVFKDISPFSY-TAPRYOT 199
QY 175 NPC--YDMKTTCLPMFGYKHVLTLDQVTRFNEEVKQSVSRNRDAPBEGFDALMOATVC 232
Db 200 NPIGKALFPCNVCYSGFGRHLPLTDVDSFNEEVKQSVSRNRDAPBEGFDALMOATVC 259
QY 233 DEKIGMNDASHLVFTTQKTHIALDGRAGIYOPNDQCHGSDNNH SASFTMDYPSL 292
Db 260 KEXIGMKRDLHLVFTTQKTHIALDGRAGIYOPNDQCHGSDNNH SASFTMDYPSL 319
QY 293 GLMTEKLSOKNINLFAVENNVNLYONYSELIPGTVGLMSDSNVQLIYDAYKIR 352
Db 320 ALLGKLAEMNINLFAVENNVNLYONYSELIPGTVGLMSDSNVQLIYDAYKIR 379
QY 353 SKVELEVRDLPEELSLFNAATCLNNEVYDGLKSCMLKIGDVSFSEIAKVRGP-OEKE 411
Db 380 SKVELEVRDLPEELSLFNAATCLNNEVYDGLKSCMLKIGDVSFSEIAKVRGP-OEKE 439
QY 412 KSTIKPVGKDSLIVQVTFDDCCACQAOAEPNSHRCNNGNGTFEGVGRGCGPGLGSC 471
Db 440 HVALRVRGGRDLEVGVIYNTGCSVGLPEPSARC-NGSGTYVGLCEGSPGLGTCR 498
QY 472 ECSEEDYRPSQDECSFREGQPVCSQREGCLGQCVCHSSDFGKIGKYCECDPFCVRY 531
Db 499 ECQDGENQSYVQNLCEABECPICSGRGDCSCNOCSEFESEFKIYGPCECDNFSCARN 558
QY 532 KGEKSGHGGSCGDCDLSQDMTGYCNCRTTRDTQMSNGLLSGRGKCEGSCVCIOP 591
Db 559 KGVLCSGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGG 618
QY 592 GSGVDTCERKPTCPDCTFKECEVECKKFPDGLHDENTCNRCROEIES-VKELMDTK 650
Db 619 GAGGEMCEKPTCPDCTFKECEVECKKFPDGLHDENTCNRCROEIES-VKELMDTK 677
QY 651 DAVNCTKNNDDCVARQYVEDSSKSIIVVEEPCPKPDILVLLVSMGAILLIGLA 710
Db 678 EAVLCFKTKAKDCVMAFTYELPSGSKNLTFLVLAPEEGNTFMAATLILAVGSLILVGLA 737
QY 711 ALLIWKLLITHDRKEFAKFEERARAKMDTANPLYKATSTFT-----NITVREG 762
Db 738 LLAIWKLIVTHDRKEFAKFEERARAKMDTANPLYKATSTFT-----NITVREG 797

RESULT 6
; Sequence 27, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Dean
; APPLICANT: Quaranta, Vito
; TITLE OF INVENTION: A NO. 5962643el Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92132
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07728, 215
; FILING DATE: 19910711
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815

```

```

; REFERENCE/DOCKET NUMBER: P31 8717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-728-215-27

Query Match 51.98; Score 2156; DB 2; Length 788;
Best Local Similarity 50.68; Pred. No. 2,1e-155;
Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

QY 1 GNICITR-----GVSSQCCIAVSPMCAMCSDEAL-PLG-SIPRCDLKENILKDNCAPE 52
Db 14 GRNDSRTNWLCLGAEETEDCLLIGPOCAMCAQENFTPSGVEKCDTPANLAKGCOLN 73
QY 53 STEFVSEARVLEDRPLSDKSGSSQVTVQSPORIALRLAPDSSKNFSIOVROVEDYPV 112
Db 74 FIENPVQVEILKKNPLSVGRKQNSSDIYQIAPOSILKLRPGCAQOTQVHVROEDYPV 133
QY 113 DLYYLMDSLMSKDDLDNIRSLGTLAEEMKRLTNSFRLGFGSVFKDISPFSY-TAPRYOT 172
Db 134 DLYYLMDSLMSKDDLDNIRSLGTLAEEMKRLTNSFRLGFGSVFKDISPFSY-TAPRYOT 192
QY 173 LENPCYDMKTTCLPMFGYKHVLTLDQVTRFNEEVKQSVSRNRDAPBEGFDALMOATVC 232
Db 193 TANPCSSIFELPLPFGRKHILPLTNDAREFRFELYKRNKISANDTPRGGCDALMOATVC 252
QY 233 DEKIGMNDASHLVFTTQKTHIALDGRAGIYOPNDQCHGSDNNH SASFTMDYPSL 292
Db 253 KEXIGMKRDLHLVFTTQKTHIALDGRAGIYOPNDQCHGSDNNH SASFTMDYPSL 312
QY 293 GLMTEKLSOKNINLFAVENNVNLYONYSELIPGTVGLMSDSNVQLIYDAYKIR 352
Db 313 GLMTEKLSOKNINLFAVENNVNLYONYSELIPGTVGLMSDSNVQLIYDAYKIR 372
QY 353 SKVELEVRDLPEELSLFNAATCLNNEVYDGLKSCMLKIGDVSFSEIAKVRGP-OEKE 411
Db 373 SEVELEVRDLPEELSLFNAATCLNNEVYDGLKSCMLKIGDVSFSEIAKVRGP-OEKE 439
QY 412 KSTIKPVGKDSLIVQVTFDDCCACQAOAEPNSHRCNNGNGTFEGVGRGCGPGLGSC 471
Db 432 KSTIKPVGKDSLIVQVTFDDCCACQAOAEPNSHRCNNGNGTFEGVGRGCGPGLGSC 491
QY 472 ECSEEDYRPSQDECSFREGQPVCSQREGCLGQCVCHSSDFGKIGKYCECDPFCVRY 532
Db 492 ECSEEDYRPSQDECSFREGQPVCSQREGCLGQCVCHSSDFGKIGKYCECDPFCVRY 548
QY 532 KGEKSGHGGSCGDCDLSQDMTGYCNCRTTRDTQMSNGLLSGRGKCEGSCVCIOP 592
Db 549 KGEKSGHGGSCGDCDLSQDMTGYCNCRTTRDTQMSNGLLSGRGKCEGSCVCIOP 608
QY 592 GSGVDTCERKPTCPDCTFKECEVECKKFPDGLHDENTCNRCROEIES-VKELMDTK 650
Db 609 GSGVDTCERKPTCPDCTFKECEVECKKFPDGLHDENTCNRCROEIES-VKELMDTK 666
QY 652 AVNCTYKNNDDCVARQYVEDSSKSIIVVEEPCPKPDILVLLVSMGAILLIGLA 711
Db 667 AVNCTYKNNDDCVARQYVEDSSKSIIVVEEPCPKPDILVLLVSMGAILLIGLA 726
QY 711 LLAIWKLIVTHDRKEFAKFEERARAKMDTANPLYKATSTFT-----NITVREG 762
Db 727 LLAIWKLIVTHDRKEFAKFEERARAKMDTANPLYKATSTFT-----NITVREG 797

RESULT 7
; Sequence 27, Application US/08938085A

```

```

: Patent No. 6339148
: GENERAL INFORMATION:
: APPLICANT: Sheppard, Dean
: APPLICANT: Quaranta, Vito
: APPLICANT: Pytela, Robert
: TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
: NUMBER OF INVENTIONS: 2
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESS: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/938,085A
: FILING DATE: 26-Sep-1997
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/728,215
: FILING DATE: 11-JUL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Parent, Annette S.
: REGISTRATION NUMBER: 42,058
: REFERENCE/DOCKET NUMBER: 023070-08021005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 27:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 788 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-938-085A-27

```

```

Query Match 51.9%; Score 2156; DB 4; Length 788;
Best Local Similarity 50.6%; Pred. No. 2.1e-165;
Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

QY 1 GPNICTR-----GVSSCOCLAVSPMCMSDEAL--PLG--SPRDLKENLTKNCAP 52
DB 14 GRNDSRTRMLCLGGAETCEDCLIGQCAMCAQENFTPHSGVGERCDTPANILAKCQ 73
QY 53 SIEFPVSEARVLDRPLSDKSGSSQVQVSPORIALRLRPDDSKNFSTIOYRVEDYV 112
DB 74 FIENPVSOYEILKNRKLPSVGRQKSSDIQIAQSLILKLRGGAOTLOVHRQETEDYV 133
QY 113 DIYVLDLSYKMDLSTIONLTKLATQMRKLTSLNRIGFADFVDPKSPMYTSPPEA 172
DB 134 DLYVLDLSYKMDLSTIONLTKLATQMRKLTSLNRIGFADFVDPKSPMYTSPPEA 192
QY 173 LBNPCVDMKTTCLPMFGYKHLVLTQVYRNEEYAKQSVSRKRDAPREGFALIMQATYC 232
DB 193 IANPCSSIPFCLPTFGFKHILPLTNDARFMEIYKNOKISANIDPEGGFALIMQAAVC 252
QY 223 DEKIGRNDASHLLVFTTAKTHIALDGRLAGIYVQNDGCHVGSNDHVSASTMDYPSL 292
DB 253 KEKIGRNDASHLLVFTTAKTHIALDGRLAGIYVQNDGCHVGSNDHVSASTMDYPSL 312
QY 293 GLATEKLSQKNINLIFAVENVVNYQYSELIPGTVGVLSMDSNNVQLITVDYAGKTR 352
DB 313 GOLIDKLVQNNVLLIFAVYQEOVHLYENYAKLIPGTVGVLSMDSNNVQLITVDYAGKTR 372
QY 353 SKVELEVRDLPELISFNATCLANNEVYIGLKSCKGLKIGTVSVSIEAKVGCPOEKRC 412

```

```

DB 373 SEVELEVLGDTGEGNLNLFPTAICNNGLTFLOHOKKCSHMKVGDIAFSYVNIPIHC-ERRSR 431
QY 413 SFTIKPVGFKDSLIVQVFPEDDCAOAAEPHSHRCNNGNGFEFGVRCRCPGMLSGCE 472
DB 432 HIIRKPVGGLDELIVSEPCNDCQKEVEVSSCHHGNSFCQGVACAPGHGMPRCE 491
QY 473 CSEEDYRPSQODECSPREQPVCSORGELGQGVCHSDPDKITGKICECDDEFCVARYK 532
DB 492 CGED---MISTSCKEAPDPHPCSGRSGRCYCGQCCILSPIGNITGPIYCGQDNFSCVARK 548
QY 533 GEMCSGHQSCGCDLCSDWGTGYCNCPTRTDTOMSSNGLICSGRGKCEGSCVCIQPG 592
DB 549 GLICGNGDCDCGECVCHSGRWGTEYCNCTSTDSVSHDGLCSRGDPCVCGKCVCTNPG 608
QY 593 SYGDCCEKCPGCPDACCTFKKECECKKDRGLAHDEMCNRYGROELISVKEIMDKDGD- 651
DB 609 ASGPTCEKPCPTGDCPCNSKRSCTIECHLSNAGAGEE--CVDKCLAGATITSEEDFSGDG 666
QY 652 AVNCTYKNEEDCVRFQYEDSSGSKILYVEEPECPGPDILVLLSVMAIILIGLAA 711
DB 667 SVSCSLGEGNECLIFLTITDNECKTIHSINEXDCPRPNIMIMLGVSLATILIGVVL 726
QY 712 LIIRKLTITIHDKFPAKFEERARAKMDTANNPLYKATISTFTNITYR 760
DB 727 ICIWKLVSHFDRKEVAKFEARSKAKWGTGNPLYRSTSTFTNNVTVK 775

```

RESULT

8

US-07-728-215-30

Sequence 30, Application US/07728215

Patent No. 5962643

GENERAL INFORMATION:

APPLICANT: Sheppard, Dean

APPLICANT: Quaranta, Vito

APPLICANT: Pytela, Robert

TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESS: Pretty, Schroeder, Brueggemann & Clark

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States of America

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/728,215

FILING DATE: 19910711

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P31 8717

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 798 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

US-07-728-215-30

```

Query Match 44.1%; Score 1831; DB 2; Length 798;
Best Local Similarity 44.9%; Pred. No. 3.3e-139;
Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

```


Db 620 CTPKFGQGTCEMCGTCTLGCACBKECVCCRAFNKEKND--TCTOECSTFNITKVESRD 677
 QY 644 EL-KDTGKDAV-NCTYKNEDEDVVRFOYEDSSGKSLIYVEEPEKPGKPDILVLLSV 701
 Db 678 KLPQVPDPVPSHCKEKEDVDCCMFYFYSVGNNEVAVHVENPECTGPDILPIYAGV 737
 QY 702 GAILLIGLALLIMKLLITTHDRKEFAKFEERAKWDTANNPLYKEATSTFTNTYRG 761
 Db 738 AGVILGLALLIMKLLIMTHDRREFAKFEKEMNAKMDTGENPIKSAVTIVVNRKYS 797

RESULT 10 US-07-728-215-29

; Sequence 29, Application US/07728215
 ; Patent No. 5962643
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/728, 215
 ; FILING DATE: 19910711
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P31 8717
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 577 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-728-215-29

Query Match 40.3%; Score 1676; DB 2; Length 577;
 Best Local Similarity 51.1%; Pred. No. 6,5e-127;
 Matches 297; Conservative 102; Mismatches 174; Indels 8; Gaps 5;

QY 121 SYSMKDDMSIONLGTAKTQKRLTSNRIFGFAVDKPVSPYATISPEELNPPCDIM 180
 Db 1 SASMDLDTLTIKELGSLSKEMSKLTSNRIIGFSGFVEKVPSPFK--TTPPEIAMPCCSI 59
 QY 181 KTTCLPMFGYKHLTLDVTRFNEEYKOSVSRNRDAPEGGFDALMOATVDEKIGRN 240
 Db 60 PYICPTPEFKHILPLTNAERFNEIYKOKISANIDNPEGGFDALMOAAYCKEKIGRN 119
 QY 241 DASHLLVFTTDAKTHIALDGLAGIVQPNQCHVSDNHSASTTMDYPSLGLMTERKLS 300
 Db 120 DSHLLVFSVSDSHSGMDSKLAGIYIPNDGLCHLDSKNEYSMSTVMEYPTIGLIDKVV 179
 QY 301 QKNINILPAVTENVVNLQNIYSELIPGTTVGVLSMDSSVNLQIYDAVGKIRSKYLEVR 360

Db 180 QNNVLLIFAVTQEQVPLXYENYAKLIPGATVGLHKKDSNNIQLIISAVEELSEVELEVL 239
 QY 361 DLPEELISFNATCLNNEVILPGLSKMGKIGDPTFSLEAKVRCCPEKRSFTIRPVG 420
 Db 240 GTEGLINTSEFSAVCCNNGTLPFHQKCLHMKVGETASFVTVSIPHC--ERKSRHYIKRVG 298
 QY 421 FSDSLIVQVTFPCDACAQAEPNHRNNGNGTEFCVCVRCGPMLSQCESEEDYRP 480
 Db 299 LBDTLIELVSPSCSDCKEYEVNSKCHNGSGVQGVACNNGHMOPHECEGDTL-- 356
 QY 481 SQDDCSPREGQPVCSQRECLCGQCVCHSSDFEKTITKYECDDFSCVRYKMGKSGHG 540
 Db 357 -STDCKETPDHPSCSGRGDVCYGCQICHLSPYGNIVGPYCCDNFSCVRHKGILCGDN 415
 QY 541 QCSGDCICDSDPMVGYCNCTTRDTCSSNGLCSGKKECCSGVCVIGQSGYDTC 600
 Db 416 DECEGCVCRSGMTGTCNCTTSDTCTISEGDTLCSGGRGDCVCGKVCVTNAGAGPTCE 475
 QY 601 CTPCDCTPFKECKEKKFDRGALHDENFCNRCRDEIESVKELKDTGKD--AVNCTYKN 659
 Db 476 CPTCSDPCNSKRSCTIECHLSADQPGGE--CYDKCKLAGVYISKADPFSKDSVSCSIQG 533
 QY 660 EDDCVRFQYEDSSGKSLIYVEEPEKPGPDILVLLSV 700
 Db 534 ENECLITFLISTDNEGKTIHINISEKDCPKPIMPIMLGV 574

RESULT 11 US-08-938-085A-29

; Sequence 29, Application US/08938085A
 ; Patent No. 6339148
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/938, 085A
 ; FILING DATE: 26-SEP-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/728,215
 ; FILING DATE: 11-JUL-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parent, Annette S.
 ; REGISTRATION NUMBER: 42,058
 ; REFERENCE/DOCKET NUMBER: 023070-0802100S
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 577 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-938-085A-29

Query Match 40.3%; Score 1676; DB 4; Length 577;
 Best Local Similarity 51.1%; Pred. No. 6.5e-127;
 Matches 297; Conservative 102; Mismatches 174; Indels 8; Gaps 5;

QY 121 SYSKMDLSIONLGTAKTARKLTSNRIFGAFVYKPVYATISPEALNCPDLM 180
 DB 1 SASMDLWTIKELGSLKSKSLSNRLFGSFEVPEVSPFK-TTPEELANPSSI 59
 QY 181 KTCLEPMFGKHLVLTDOVTRNEEYKQSVSRNDAPEGFDIMQATVDEKIGMN 240
 DB 60 PYICLTFEGFKHLPLTNAERFNEVYKOKISANDNEGGFDALMOAAVCEKIGMN 119
 QY 241 DASHLLVETTDKTHALDGLAGIVQPDGQCHVSDNHYASTTMDYPSLGLMEXIS 300
 DB 120 DSHLLVETTDKTHALDGLAGIVQPDGQCHVSDNHYASTTMDYPSLGLMEXIS 179
 QY 301 OKINILFVTEVNVVLYXONSLLPGTAVGLMSDSSVNLQIYAVGKIRSKYLEVR 360
 DB 180 QNVNLLIFAVTOEPLVYEVAKLIPGATVGLHDKSNGILQILISAVELKSEVLEVL 239
 QY 361 DLPEELSFNATCLNNEVYIPGLKSCMLKIGDVSFSIEAKVRCGPQEKESFTIKPVG 420
 DB 240 GDEGLNLSFSAVNCNGTJLPHQKCLHMKVGETASFNTVSIPLNC-EKSRHVILKPVG 298
 QY 421 FKDSLIVQVTPDCACQAOAEPNSHRCNNGCTFECGYCRCGPGMLGSCQCESEEDYRP 480
 DB 299 LGDTLILVSPESCSCCQKQEVENSSKCHNGNSYOCACNPGHMGPHCEGEBTTL-- 356
 QY 481 SQODESPRAGOVPSORGCLOGCVCVSHSDGKITGKVECDDESCVRYGEMSGHG 540
 DB 357 -STDCKETPDHPSGSGRGCTCGCICHLSPGNITGYCQCDNCSVRHNGLLGCDNG 415
 QY 541 QCSGDLCDSDMTGYCCNCTRTDQMSNGLCSGRKCEGSCVCIOPSGYDGTCPK 600
 DB 416 DCGEGVCSGWTGXECNCTSTPCISPDGTLCSGRGCVCGKVCNPGASGPTCR 475
 QY 601 CPTCPDCTKKECECKEKFDRGALDENTCNRYCDELESVELKDTGKD-AVNTYKN 659
 DB 476 CPTCSPCNKSRKICHSADQPGEE--CYDKCLAGVTLISKADFSSVSCSLQG 533
 QY 660 EDDCVAFQYEDDSGKSLVYVEEPCPKPDLVLLSV 700
 DB 534 ENCLITFLSTNEGTITIHNSKDCPKRPIMIMIGV 574

RESULT 12
 US-07-728-215-33
 : Sequence 33, Application US/07728215
 : Patent No. 5962643
 : GENERAL INFORMATION:
 : APPLICANT: Sheppard, Dean
 : APPLICANT: Quaranta, Vito
 : APPLICANT: Pytela, Robert
 : TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 : NUMBER OF SEQUENCES: 43
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 : STREET: 4370 La Jolla Village Drive, Suite 700
 : CITY: San Diego
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 92122
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/728, 215
 : FILING DATE: 19910711
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P31 8717
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-8949
 TELEFAX: (619) 535-9001
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 846 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear

US-07-728-215-33

Query Match 37.0%; Score 1535.5; DB 2; Length 846;
 Best Local Similarity 38.7%; Pred. No. 2.3e-115;
 Matches 320; Conservative 128; Mismatches 292; Indels 87; Gaps 21;

QY 4 ICTTGVSSCOQCLAVSPMCASDEALPLGSPDCLEKLNKNCAPESTIEPPVSEARV 63
 DB 37 VSTCAKSKCHTCTQTEG-CAMCQDPFK-GQSRCYQNTSSL--CFEEFAVSPITVEOI 91
 QY 64 LEDRPL-----SDKSGDSS----- 78
 DB 92 LVNKNLTQVYAEELAGGGGAGSGSSSSSYSSSSSSSFYSQSSGSSASGYEYSAG 151
 QY 79 QVTQVSPORIALRLPDDSKNFSIOVRVEDYPVDIYVLMDSYSMKDLSIONLGTKL 138
 DB 152 EIVQVQPSMLALRVNKNHNIKISYQAGEYVPDYLVLMDSYSMEDKAKLSTLDKLT 211
 QY 139 ATQKRLTSLNLRIGFAGVYKPVSPYATISPEALNCPYMKTKTCLPMFGKHLVLT 198
 DB 212 SETMRITNNHLLGSGSVFDKVLMPYV-STIPKLEHPC---ENCKAPYQVQNHMLPN 266
 QY 199 QVTRFNEEYKQSVSRNDAPEGFDALMOATVCDERIGMENDSHLVTFTDAKTHAL 258
 DB 267 NTESFSNEKNAVYSGNLDAPEGFDALMOAIAHRSQIGRHEQARLLVFTSDGPHAG 326
 QY 229 DRLAGIVQPDGQCHVSDNHYASTTMDYPSLGLMEXISQNNINILFVTEVNVLY 318
 DB 327 DGLGGLVAPNDGEBCHLPKGEYHSTLDYPSISQINQKVDNALINILFVAVTASOLSV 386
 QY 319 QNVSELIGTAVGLMSDSSVNLQIYAVGKIRSKYLEVRDLPEELSFNATCLNNE 378
 DB 387 EKLVEHIGSSSAKLDNDSSNVLELKEEYKRISSVEMK-DNATGDKITTFSSCLSN 445
 QY 379 VTPGLKSCMLKIGDVSFSIEAKVRCGPQEKESFTIKPVGFKDSLIVQVTPDCD 435
 DB 446 PEVQTSKCDNLKEGOVGFYTAQIOLKCPEDRDMTORIHISPVGINVMDIOLMLCSC 505
 QY 436 ACQ-----AQAEPNSHRCNNGNGTTECGVCRGFGMLGSCQCESEEDYRPSODESPR 489
 DB 506 PCENSGSIGYQVANS--C-SGHGTSKMGICNCDDSYGNGCECATLDTLSKFNANDSCR 562
 QY 490 EGQPV---CSQGBELCQCVCHS--SDFGKITGKVECDDESCVRYGEMSGHG--HGQC 542
 DB 563 ADSTTDCSGGHCHCVACCECHKRPNPLEITISGHCHCDNDSCKRNNOQLCSGDHGTG 622
 QY 543 SCGDCLCDSDMTGYCCNCTRTDQMSNGLCSGRKCEGSCVCI--QPSGYDGTG 598
 DB 623 ECGRCKCKRGWTSNGCGQESNDTCLMPGGGEGICSGHGTCEGVCYKCVINOGRSSGHC 682
 QY 599 EKCPTCPDCTKKECECKEKFDRGALDENTCNRYCDE---ETESVELKDTGKDAVN 654
 DB 683 EKCPTCSGRCLKDCVOCOMYKTGELKNGDCARNCQFVGVGEVY-EIDET-KDEQM 740
 QY 655 CTYKNEDECVAVFQYEDDSGKSLVYVEEPCPKPDLVLLSVKMAIILIGLAALI 714
 DB 741 CFEEDEDEKEMFKYSE--QGLHAYAOENKCEPKAVVLMGLGVIAIYVGLAILL 798
 QY 715 WKLLTITHDKKEFAFEERAKAKMDTANNPYKATSTFTNITYRG 761

Db 799 WKLLTTHDRREFAREKERMAKMDTGENDPIYKQATSTFKNPMYAG 845

RESULT 13

US-08-938-085A-33
Sequence 33, Application US/08938085A

Patent No. 6339148

GENERAL INFORMATION:

APPLICANT: Sheppard, Dean

APPLICANT: Quaranta, Vito

APPLICANT: Pyeala, Robert

TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses

TITLE OF INVENTION: Theoreof

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/938,085A

FILING DATE: 26-SEP-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/77,281,215

FILING DATE: 11-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Parent, Annette S.

REGISTRATION NUMBER: 42,058

REFERENCE/DOCKET NUMBER: 023070-080210US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 846 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-938-085A-33

Query Match 37.0%; Score 1535.5; DB 4; Length 846;

Best Local Similarity 38.7%; Pred. No. 2.3e-115;

Matches 320; Conservative 128; Mismatches 292; Indels 87; Gaps 21;

Db 4 ICTTRGVSSCCQCLAVSPWCAMCSDALPLIGSPRCDLKENLLKDKCAEESIEFFPSEARV 63

Db 37 VTCSKSKKCHTCIOETG-CAMCQMPDCK-GQSRCTYQNTSSL--CPEFAFYSPTIVBOI 91

Db 64 LEDRPL-----SDKSGDSS-----78

Db 92 LVNNKLTNQYKALAGGGGAMSS 151

Db 79 QVTQVSPQIALRLRPDSKNFSIYQVQVEDYPVDIYLLMDISYMKDLSIONLGLTKL 138

Db 152 EIVQIOPQSMRLALRVNEKHNIKISYQAEQYVDLYLLMDLSKSMEDDKAKLSTLGDKL 211

Db 139 ATQMRKLSNLIRIGGAFQDKVSPYMYISPEALENFCYDMKTKCLMFGYKHYLTLD 198

Db 212 STTKRITNNFALFGSEFQDKLMPYV-STIPKLEHPC---ENCKAPGYQNMPLNN 266

Db 199 QVTRNEEVKQSVARNDAPEGFDALMOATVCDKIGMRNDASHLLVFTTDAKTHAL 258

Db 267 NTESFSNEKNAIVTSQNDAPBGGFDALMOALACNSQIGMRQARRLVFTSDAGFHAG 326

Qy 259 DGRLAGIOPNDGCHVSDNHHYSASTMDYPSIGLMTKESLOKNNILFEAVTENVNL 318

Db 327 DKLGLGVIAIPNDGCHLSPKGEYTHSTLDYPSISQINQKVDKNAINIFAVTASQSLV 386

Qy 319 QNYSILIPGTTVGVLSMDSNVQLQIDVADYKISKVLELEVRDLPPELSTLSFNATCLNE 378

Db 387 EKLVEHIGSSAAKLDNDSSNVVELKREYKRISSVEMK-DNATGVKITYFSSCSLNG 445

Qy 379 VTPGLKSCMGKIGDVTGFSFISDAKVRGCPQERE--KSFITKPVGEKSLIVQVTEPDC 435

Db 446 PEVQTSKCDNLKEGQVAFQAIQILKCPEDPRMTQTIHISPVGINEVMOIQLTMLCSC 505

Qy 426 ACQ-----AQAPSHHCNNGNTEFEGYVRCRPGMIGSCQSESEEDYRSDQDECS 489

Db 506 PCENPQSGIYQVQANS--C-SGHGYSKICGICNDQSYFGNKECSAIDLKSPANDSCR 562

Qy 490 EGQPV---CSQRGELCQGVCHS--SPFGKITGKCECDPFCVRYKGEKCSG--HGQC 542

Db 563 ADSTSTDCSGRGHCVACACHRPNIELISGKHCECDNFSCERRNRNQLSGPDHGT 622

Qy 543 SCGDCLCDSDMTGYCNCCTTPTICM--SSNGLCSGRKCEGSCVCI---QPGSYGDT 598

Db 623 ECGRCKCKPMTGSGNCGQESNDTCMPGGGEGISGHTGCEGVCCTVNDQGRFSGRHC 682

Qy 599 EKCPCPACTFEKKECVCEKFEKFDGALHDEMTGRCYRD---ETESVKELKDTGKDAVN 654

Db 683 EKCPCSCRCGLKDCVQCCQMYKTELKNGDCAKNTQYVPVGEVY-ETDET-KDEQM 740

Qy 655 CTYKNEDDCVVRFQYEDSSGKSLIYVEEPCPKGPDILVLLSVGAILLIGLALLI 714

Db 741 CKFEDEDCRFMFYKSE--QELHVVYAQENNECPAKFVGLGVVIAIVLGLATLL 798

Qy 715 WKLLTTHDRKPEFAKFEERARAKMDTANPLLYKEASTFTNTYRG 761

Db 799 WKLLTTHDRREFAREKERMAKMDTGENDPIYKQATSTFKNPMYAG 845

RESULT 14

US-08-789-078-1
Sequence 1, Application US/08789078

Patent No. 5843885

GENERAL INFORMATION:

APPLICANT: Benedict, Stephen H.

APPLICANT: Siahaan, Teruna

APPLICANT: Chan, Marcia

APPLICANT: Tibbets, Scott

TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING

TITLE OF INVENTION: IMMUNE TOLERANCE

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS

STREET: 1101 Walnut St.

CITY: Kansas City

STATE: MO

COUNTRY: USA

ZIP: 64106

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/789,078

FILING DATE: 03-FEB-1997

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/229,513

FILING DATE: 19-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Collins, John M.

REGISTRATION NUMBER: 26262

REFERENCE/DOCKET NUMBER: 22833


```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816)474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label= "signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 449..496
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label= "repeat"
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label= "trans"
OTHER INFORMATION: /note= "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label= "cyto"
OTHER INFORMATION: /note= "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TITLE: LFA-1 Amino acid sequence (82) (from human
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
US-08-789-078-1

```

```

Query Match 36.4%; Score 1513.5; DB 2; Length 769;
Best Local Similarity 39.9%; Pred. No. 1,2e-113;
Matches 308; Conservative 133; Mismatches 278; Indels 53; Gaps 21;

```

```

OY 119 DISYKDDLSINQNGTKLATOMKRLTSLNIGCAFVDKPSVPMATISPEALENFCY 178
DB 134 DLSYMLDRLNRVKKLGDLALALNEITESGRIGSVFVTLVFN-TAPDKLRNCP 192
OY 179 DMKTCLEPFGYKHYLTLLDQVTRFNEVEYKOSVARNDAPEGGDALIMQATVCEKIGM 238
DB 193 NKEEQPEFAFHVKLKLNNSNOFOTEGKOLISGNDAPGGLDAPGAAACPEEIGM 252
OY 239 RNDASHLLETTDAKTHIALDGRLAGIYVNDGQCHVSDNHSASTTMDYPSGLMTER 298
DB 253 RN-VTRLVFAFDGDFHFAQDGKLGAILLPNDGRHL-EDNLKRSNEDFPPVQQLAHK 310
OY 299 LSKNINILIPAVENVNLKQNSLIGTYGVLSMSSNVLQILVAYKIRKAYLE 358
DB 311 LAENNIOPIFAVTSRMVYKTEKLEIPKSAVGELESDSSNVVHLIKNAKLSRYFLD 370
OY 359 VRDPEELSLFNATCLNNEVIRGLK-----SCMGKIGDYFSFEAKVRGCPQKE 411
DB 371 HNALPDTLKVTYDSFCN-----GYTHRNQPRGDDCGQINPITFOVKVATTECIG--E 423
OY 412 KSFTRKPVGFKDSLIVQVTFDCACQAOAEPNSHRCNNGNTEFCGYCRGPGNLSQC 471
DB 424 QSFVIRALGETDIYVQVLPQCECRDRS-RDLS-LG-HGKFLGALRCDDTYIGKNC 481
OY 472 ECSEEDYRPSQDE--CSFREGQPVCSQREGELCGQCYCHSDF-GK-ITGKYCCDPS 527
DB 482 ECOTOG-RSSQELBGSCKRDNNSIICSGLDGCVCCQCLHSDVPGKLIYGYCCDDIN 540
OY 528 CVRYKEMSGHQ--CSCGDCDLDSDMTGYCNCCTRTDTCMSNGLLCSGRGCEGS 585
DB 541 CERNGQVCGGFGRLGFCGKCRCHGFGSCQCEBRTTEGLNRRRECSRGRCNV 600
OY 586 CVCIOPGSYGDCCEKQPCPCDCTFKKECECKKDRGALHDEMCNRYCNDEISVREL 645
DB 601 CEC-HSGYQLPLCQCEPCGSPCGKYISCAECLKEKPF--GKNCSAACPG-----LQL 652
OY 646 KDTGKDAVNCYTKNEDDCYVRFQY-YEDSSGKILYVVEPCPGPOLYVLLSVMAI 704
DB 653 SNNPVKRGTRKEDSGCAVATLEQDDMDRYLLYVESRECVAGPNALAVGTVAGI 712
OY 705 LILGLALLIMKLLITIHDKREFAFEEERARAKMDTANPNLYKATSTFTN 756
DB 713 VLIGILLIYIMKALHLSDLREYRREKELKSQNN-DNPLFKSATITVNM 763

```

```

RESULT 15
US-08-752-633-1
Sequence 1, Application US/08752633
Patent No. 5863889
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Slahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbets, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,633
FILING DATE:
CLASSIFICATION: 514

```

ATTORNEY/AGENT INFORMATION:
 NAME: Collins, John M.
 REGISTRATION NUMBER: 26262
 REFERENCE/DOCKET NUMBER: 22833
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (816)474-9050
 TELEFAX: (816)474-9057
 TELEX: 434-363
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 TOPOLOGY: both
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 TISSUE TYPE: tonsil
 FEATURE:
 NAME/KEY: Region
 LOCATION: 1..22
 OTHER INFORMATION: /label= signal
 FEATURE:
 NAME/KEY: Region
 LOCATION: 449..496
 OTHER INFORMATION: /label= repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 541..581
 OTHER INFORMATION: /note= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 497..540
 OTHER INFORMATION: /label= repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 582..617
 OTHER INFORMATION: /note= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 701..723
 OTHER INFORMATION: /label= trans
 OTHER INFORMATION: /note= "transmembrane domain"
 NAME/KEY: Domain
 LOCATION: 724..769
 OTHER INFORMATION: /label= cyto
 OTHER INFORMATION: /note= "cytoplasmic domain"
 PUBLICATION INFORMATION:
 AUTHORS: Pigott,
 TITLES: LFA-1 Amino acid sequence (B2) (from human
 TITLE: tonsil)
 JOURNAL: The Adhesion Molecule Facts Book
 PAGES: 96-96
 DATE: 1993
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
 US-08-752-633-1

Query Match 36.4%; Score 1513.5; DB 2; Length 769;
 Best Local Similarity 39.9%; Pred. No. 1.2e-113;
 Matches 308; Conservative 133; Mismatches 278; Indels 53; Gaps 21;

QY 5 CTRGVSSCQCCIAVSPMCAMCSDALPLGSP-----RCDLKENILKDNCAPESEIEPPV 58
 DB 25 CTRGVSSCQCCIAVSPMCAMCSDALPLGSP-----RCDLKENILKDNCAPESEIEPPV 58

QY 59 SEARVLEDRPLSDKSGSDSSQVTVSQRALRLRPDSDSKNESTQVROVEDYPIVYIM 118
 DB 83 SLAEIQEDHNGGK-----QLSPQKVTLYLRPGGAALFNWTFRRAKYPIIDLTYIM 133
 QY 119 DLSYMKDLDLWISQIMGLKATQMRKLTNLRIGAFVDRPVSYMISPPALENCPY 178
 DB 134 DLSYMKDLDLWISQIMGLKATQMRKLTNLRIGAFVDRPVSYMISPPALENCPY 192
 QY 179 DMKTTCLPMFGYKHVLTITDOVTRNEEYKQSVSRNDAPEGGDAIMQATVCEKIGW 238
 DB 193 NKEKECPPEAFRHYLKLITNNSNQFTEVGKQLISGNLDAPEGIDAMQVAAPPEEIGW 252
 QY 239 RNDASHLVETTDAKTHIALDGRLAGIVQPDGCHGSDNHYASATJMDVPSGLMTER 298
 DB 253 RN-VTRLVFATDGFHFRAGCKLGAITPMDGRCHL-EDNLYKRSNFEIPSGQLAHK 310
 QY 299 LSQKNINLIFAVTENVMVLYONSELIPTGVVLSMDSNVLIQIYDAVKIRKSYELE 358
 DB 311 LAENNIPFVATSMVTKYETKLEIPKSAVGELEDSDSNVHLIKNAVKLSSRVFLD 370
 QY 359 VRDLPPEELSPFNATCLNNEVIPLGK-----SCMKLIGDIYSPESIENKVRGCPPEKE 411
 DB 371 HNALPDTLKVYIDSPCSN-----GVTRNRPDGDGDGVQINWPTFQVKYATECIQ--E 423
 QY 412 KSFITKPYGFXDSLIVQYTFDCCACQAQAPNSHRCNNGTTECGVCRCGPGMLGSCQ 471
 DB 424 QSEVIRALGFTDIYVQVLPQCECRQRQSDRS-LC-HKGFLGEGICRDTYIGKNC 481
 QY 472 ECSEDIYPSQODE--CSPREGQVYCSRGELCGQCYCHSSD-GR-ITGKYCECDFE 527
 DB 482 ECQYQG-RSSQELGSGCRKDNNSIICSLGICGCGCLCHSDVPRKLIQYCECDPTIN 540
 QY 528 CVRYKGEKSGHGO--CSCGDLCDSDMTGYCCTRTDPCMSNGLCSGRKCEGS 585
 DB 541 CERNMGVCGSPGRGLCFCCGKCHPGEFSGACCCERTTEGCLNPRVYECSSGRRCNCV 600
 QY 586 CVCIOPGSYGDTCEKCPDPACTCFKKECYECKKFDGALHDEYCNRYCRDEIESYKEL 645
 DB 601 CEC-HSGYQLPLCQCEPCPCSPCKYISACCLFEKGP--GRKCSAACPG-----LQL 652
 QY 646 KDTGKDAVNCYKKNEDCVARFQY-YEDSSKSLIYVEEPECKGDLIVLLSVGAI 704
 DB 653 SNPNYKGTCKERSESGGWAAYTLEQDDGMRYLIYDESECVAGNIAIAGVTYAGI 712
 QY 705 LLIGLAALLIKLLITTHDKREFAKFEERARAKMTANNPLYEASTFTN 756
 DB 713 VLIGILLVIMKALIHLSDLREYRFRFEKIKLSQMN--DNPDKRSATTVMN 763

Search completed: May 19, 2002, 10:39:46
 Job time: 9115 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 09:42:13 ; Search time 55.89 seconds
(without alignments)
1310.075 Million cell updates/sec

Title: US-09-673-302a-1
Perfect score: 4154
Sequence: 1 GPMICITRGVSSCOCCLAVS.....NMLEYKATSTFTNTYRGT 762

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pIR.71:*
2: pIR.1:*
3: pIR.3:*
4: pIR.4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4154	100.0	788	2	platelet glycoprot
2	4146	99.8	778	2	platelet glycoprot
3	4119	99.2	788	2	platelet glycoprot
4	4038	97.2	753	2	platelet glycoprot
5	3577	86.1	723	2	integrin beta-3 ch
6	3484.5	83.9	781	2	integrin beta-3 - c
7	3470	83.5	680	2	integrin beta-3 ch
8	3260.5	78.5	788	2	integrin beta-3 su
9	2340	56.3	799	2	integrin beta-5 ch
10	2156	51.9	788	2	integrin beta-6 ch
11	1994.5	46.0	656	2	integrin beta-5 ch
12	1846.5	44.5	803	1	integrin, band 3 p
13	1831	44.1	798	2	fibronectin recept
14	1826.5	44.0	798	2	integrin beta-1* c
15	1821	43.8	798	2	integrin beta-1 ch
16	1818.5	43.8	798	2	integrin beta-1 ch
17	1817	43.7	773	2	beta-1 integrin su
18	1806.5	43.5	799	1	fibronectin recept
19	1801.5	43.4	799	1	integrin beta o1lg
20	1676	40.3	809	2	integrin beta-6 ch
21	1579	38.0	809	2	integrin beta-4 ch
22	1553	37.4	772	2	integrin beta 2 ch
23	1547.5	37.3	846	2	integrin beta ch1
24	1520.5	36.6	769	1	leukocyte adhesion
25	1517	36.5	771	2	leukocyte adhesion
26	1515	36.5	770	2	leukocyte adhesion
27	1513.5	36.4	769	1	leukocyte adhesion
28	1506	36.3	806	2	integrin beta-7 ch
29	1502.5	36.2	798	2	integrin beta-7 ch

30	1331	32.0	768	2	B41029	integrin beta-8 ch
31	1326.5	31.9	769	2	A41029	integrin beta-8 ch
32	1184	28.5	964	2	JC5545	integrin beta-4 pr
33	1184	28.5	1875	2	A36429	integrin beta-4 ch
34	1154.5	27.8	1807	2	JC6319	integrin beta-4 ch
35	1145.5	27.6	1748	1	JN0786	integrin beta-4 ch
36	948	22.8	277	2	S14324	platelet glycoprot
37	738	17.8	150	2	I70163	integrin beta 3 su
38	545	13.1	92	2	D37057	epithelial cell g1
39	419	10.1	194	2	I51310	beta 8 integrin -
40	272	6.5	95	2	S6396	integrin beta-1 ch
41	268.5	6.5	2019	1	JQ1322	tenascin precursor
42	265	6.4	2406	2	A54148	od3 protein - fru1
43	265	6.4	2515	2	S47008	tenascin-like prot
44	260.5	6.3	3566	1	A40701	tenascin-X precurs
45	259	6.2	2201	2	A32160	tenascin-C - human

ALIGNMENTS

RESULT 1
177349
platelet glycoprotein IIb precursor - human
N:Alternate names: GPIIb
C:Species: Homo sapiens (man)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 04-Mar-2000
C:Accession: I77349; F56793
R:Frachet, P.; Uzan, G.; Thevenon, D.; Denarier, E.; Prandini, M.H.; Marguerie, G.
Mol. Biol. Rep. 14, 27-33, 1990
A:Title: GPIIb and GPIIb amino acid sequences deduced from human megakaryocyte cDNAs
A:Reference number: 157461; MUID:90265363
A:Accession: I77349
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-788 <FRA>
A:Cross-references: GB:M35999; NID:q183532; PIDN:AAA35927.1; PID:q306795
R:Cailliet, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIc*, GPIIa
A:Reference number: A56793; MUID:92061944
A:Accession: F56793
A:Status: preliminary
A:Molecule type: protein
A:Residues: 27-30, 'X', 32-37 <CAN>
A:Experimental source: platelet
C:Keywords: glycoprotein; platelet
F:1-71/Domain: signal sequence #status predicted <SIG>
F:123,346,478,782/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 100.0%; Score 4154; DB 2; Length 788;
Best Local Similarity 100.0%; Pred. No. 1.5e-256;
Matches 762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GNITTTGVSSCOCCLAVSPMCAMCSDEALPLGSPKDLKLNLCNCAESTEPFSE	60
DB	27	GNITTTGVSSCOCCLAVSPMCAMCSDEALPLGSPKDLKLNLCNCAESTEPFSE	86
QY	61	AVLEDRPLSDKSGDSQVQVSPQIALRLRPDSSNFSGIQVROVEDYVVDIYLLMDL	120
DB	87	AVLEDRPLSDKSGDSQVQVSPQIALRLRPDSSNFSGIQVROVEDYVVDIYLLMDL	146
QY	121	SYSMKDLMSIONIGTKLATQWRKLTSLNRIGFAFVDKPVSPYWTSPPEALENPECYDM	180
DB	147	SYSMKDLMSIONIGTKLATQWRKLTSLNRIGFAFVDKPVSPYWTSPPEALENPECYDM	206
QY	181	KTTCPMPGKYKIVLLTQVYFRNEEVKKOSVSRNRDAPPEGFPAIMQATYCDKXIGWRN	240
DB	207	KTTCPMPGKYKIVLLTQVYFRNEEVKKOSVSRNRDAPPEGFPAIMQATYCDKXIGWRN	266
QY	241	DASHLVLFTTDAKTHIALDGRLAGIVQPNDCQCHGSDNHNYSASTTMDYPSLGLATEKLS	300

```

Db 267 DASHLVEFTDAKTHIALDRLAGIVQPNDOCHVSDNHYASTMDYPSLGLMTEKLS 326
QY 301 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 360
Db 327 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 386
QY 361 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 420
Db 387 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 446
QY 421 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPGMLGSCCESEEDYRP 480
Db 447 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPGMLGSCCESEEDYRP 506
QY 481 SQODECSPRGQPVCSQRGECLOGCVCCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 540
Db 507 SQODECSPRGQPVCSQRGECLOGCVCCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 566
QY 541 QCSGDCDCLDSMTGYCNCCTRTDTCMSNGLLCSGRKCEGSCVCIOPGSYGDTCKR 600
Db 567 QCSGDCDCLDSMTGYCNCCTRTDTCMSNGLLCSGRKCEGSCVCIOPGSYGDTCKR 626
QY 601 CPTCPDCTFKKCEVCKKFRDGLHDETCNRYCDELESYKELKDTGKDAVNCTYKNE 660
Db 627 CPTCPDCTFKKCEVCKKFRDGLHDETCNRYCDELESYKELKDTGKDAVNCTYKNE 686
QY 661 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIMKLIT 720
Db 687 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIMKLIT 746
QY 721 IHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNITYRGT 762
Db 747 IHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNITYRGT 788

RESULT 2
A:Accession: A60798
A:Molecule type: DNA
A:Residues: 1-778 <ROS>
A:Cross-references: GB:M57494
A:Title: Characterization of the human platelet glycoprotein IIa gene. Comparison with
A:Reference number: A36085; MUID:91009291
A:Accession: A60798; MUID:88294314
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-778 <ROS>
A:Cross-references: GB:M57494
A:Title: Characterization of the human platelet glycoprotein IIa gene. Comparison with
A:Reference number: A36085; MUID:91009291
A:Accession: A36085
A:Molecule type: DNA
A:Residues: 1-705, 'G', 707-778 <LAN>
A:Cross-references: GB:M57494
A:Title: Characterization of the human platelet glycoprotein IIa gene. Comparison with
A:Reference number: A36085; MUID:91009291
A:Accession: A36085
A:Molecule type: protein
A:Residues: 429-433 <HIR>
A:Title: The genomic organization of platelet glycoprotein IIa.
A:Reference number: A36268; MUID:90256778

```

```

A:Accession: A36268
A:Molecule type: DNA
A:Residues: 18-705, 'G', 707-778 <ZIM>
A:Cross-references: GB:J05427
A:Note: The authors translated the codon GAT for residue 233 as Glu, GAT for residue
C:Genetics:
A:Gene: GDB:117GB3
A:Cross-references: GDB:120013; OMIM:173470
A:Map position: 17q21.32-17q21.32
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: alternative splicing; glycoprotein

Query Match 99.8%; Score 4146; DB 2; Length 778;
Best Local Similarity 99.9%; Pred. No. 4.6e-26;
Matches 761; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNICTTGRVSSCOQCLAVSPMCAMCSDDEALPLGSPRODLKENLKNCAPESEIEPVSE 60
Db 17 GNICTTGRVSSCOQCLAVSPMCAMCSDDEALPLGSPRODLKENLKNCAPESEIEPVSE 76
QY 61 ARVLEDRPLSKGSDSSQVTVSPORIALRLRPDSSKNFSIOVQYVEDYVPVDIYMLD 120
Db 77 ARVLEDRPLSKGSDSSQVTVSPORIALRLRPDSSKNFSIOVQYVEDYVPVDIYMLD 136
QY 121 SYSMKDLMSIONLGTKLATQMRKLTSMRLRGFAFVDKPVSPMYTSPPEALENFCYDM 180
Db 137 SYSMKDLMSIONLGTKLATQMRKLTSMRLRGFAFVDKPVSPMYTSPPEALENFCYDM 196
QY 181 KTTCLPMGYKXHVLLTDQVTRFNEEYKQSVSRNRDAPGEGPDALMQAVVCEKIGWRN 240
Db 197 KTTCLPMGYKXHVLLTDQVTRFNEEYKQSVSRNRDAPGEGPDALMQAVVCEKIGWRN 256
QY 241 DASHLVEFTDAKTHIALDRLAGIVQPNDOCHVSDNHYASTMDYPSLGLMTEKLS 300
Db 257 DASHLVEFTDAKTHIALDRLAGIVQPNDOCHVSDNHYASTMDYPSLGLMTEKLS 316
QY 301 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 360
Db 317 OKNINLIFAVTENVVNLKYONSELIPGTGYVLSMDSNVLOLIVAYGIRSKVELEVR 376
QY 361 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 420
Db 377 DLPEELSLSFNATCLNNEVTPGLKSCMGLKIGDVSFSEIAKVRGCPQEKESFTIKPYG 436
QY 421 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPGMLGSCCESEEDYRP 480
Db 437 FKDSLIVQVTFDDCCACQAOAEPSNRHRCNNNGTFECGVCRCGPGMLGSCCESEEDYRP 496
QY 481 SQODECSPRGQPVCSQRGECLOGCVCCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 540
Db 497 SQODECSPRGQPVCSQRGECLOGCVCCHSSDPFKITGKYCECDDPSCYRYKGMCSGNG 556
QY 541 QCSGDCDCLDSMTGYCNCCTRTDTCMSNGLLCSGRKCEGSCVCIOPGSYGDTCKR 600
Db 557 QCSGDCDCLDSMTGYCNCCTRTDTCMSNGLLCSGRKCEGSCVCIOPGSYGDTCKR 616
QY 601 CPTCPDCTFKKCEVCKKFRDGLHDETCNRYCDELESYKELKDTGKDAVNCTYKNE 660
Db 617 CPTCPDCTFKKCEVCKKFRDGLHDETCNRYCDELESYKELKDTGKDAVNCTYKNE 676
QY 661 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIMKLIT 720
Db 677 DDCVVRFOYVEDSSGKSILYVEEPCPKGPDILVLLSVMGAILLIGLAALLIMKLIT 736
QY 721 IHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNITYRGT 762
Db 737 IHDRKEFAKFEERARAKWDTANNPLYKEATSTFTNITYRGT 778

RESULT 3
A26547
Platelet glycoprotein IIa beta chain precursor (version 1) - human

```

N:Alternate names: endothelial cell glycoprotein IIA
 C:Species: Homo sapiens (man)
 C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 20-Aug-1999
 C:Accession: A26547; A32532; S65439
 R:Flynn, L.A.; Steinberg, B.; Reil Jr., S.C.; Lo, S.; Phillips, D.R.
 J. Biol. Chem. 262: 3936-3939, 1987
 A:Title: Protein sequence of endothelial glycoprotein IIA derived from a cDNA clone. ID
 A:Reference number: A26547; MUID:87165991
 A:Accession: A26547
 A:Molecule type: mRNA
 A:Residues: 1-788 <FT>
 A:Cross-references: GB:J0703; NID:g183452; PIDN:AAA52589.1; PID:g306786
 R:Zimlin, A.B.; Elsmann, R.; Vlahire, G.; Schwartz, E.; Bennett, J.S.; Poncz, M.
 J. Clin. Invest. 81, 1470-1475, 1988
 A:Title: Structure of platelet glycoprotein IIA. A common subunit for two different men
 A:Reference number: A32532; MUID:88213696
 A:Accession: A32532
 A:Molecule type: mRNA
 A:Residues: 1-11, 'A', 13-788 <ZIM>
 A:Cross-references: GB:M20311; NID:g190107; PIDN:AAA60122.1; PID:g190108
 R:Makogonko, E.M.; Yakubenko, V.P.; Ingham, K.C.; Medved, L.V.
 Eur. J. Biochem. 237, 205-211, 1996
 A:Title: Thermal stability of individual domains in platelet glycoprotein IIBIIIA.
 A:Reference number: S65437; MUID:96203926
 A:Accession: S65439
 A:Molecule type: protein
 A:Residues: 27:375;414 <MAK>
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: alternative splicing; cytoskeleton; glycoprotein; transmembrane protein
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:17-188/Domain: signal sequence #status predicted <SIG>
 F:119-747/Domain: transmembrane #status predicted <TM>
 F:125,346,397,478,585,680,782/Binding site: carbohydrate (asn) (covalent) #status predic

Query Match 99.2% Score 4119; DB 2; Length 788;
 Best Local Similarity 99.3% Pred. No. 2, 5e-254;
 Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTTGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPPVSE 60
 DB 27 GPNICTTGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPPVSE 86
 QY 61 ARVLEDRPLSDKSGSSQVTOVSPORIALRLRPDDSKNFSIOVROYEYPPDIYLLML 120
 DB 87 ARVLEDRPLSDKSGSSQVTOVSPORIALRLRPDDSKNFSIOVROYEYPPDIYLLML 146
 QY 121 SYMKDDLMSTIONLGRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENFCYDM 180
 DB 147 SYMKDDLMSTIONLGRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENFCYDM 206
 QY 181 KTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGFDALMQATVDEKIGMNR 240
 DB 207 KTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGFDALMQATVDEKIGMNR 266
 QY 241 DASHLVFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLS 300
 DB 267 DASHLVFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLS 326
 QY 301 OKNNTLFAVTENVNLYONVSELIGETVGLSMDSNVLOLIVAYGKIRSKVELEVR 360
 DB 327 OKNNTLFAVTENVNLYONVSELIGETVGLSMDSNVLOLIVAYGKIRSKVELEVR 386
 QY 361 DLPELSLSFNATCLNNEVTPGLSKCMGLRTGTVSFSIAKVRGPOEKESFTTKPVG 420
 DB 387 DLPELSLSFNATCLNNEVTPGLSKCMGLRTGTVSFSIAKVRGPOEKESFTTKPVG 446
 QY 421 FKDSLIVQVTFDDCCACQAOAEPNSHRCCNNGNTEFEGVCRGPGWGLSGCCESEEDYRP 480
 DB 447 FKDSLIVQVTFDDCCACQAOAEPNSHRCCNNGNTEFEGVCRGPGWGLSGCCESEEDYRP 506
 QY 481 SQODEGSPRGOPVCSQRGBCLGQGCCHSSDREKTKGKCEDDPFCVYKQEMSGHG 540

DB 507 SQODEGSPRGOPVCSQRGBCLGQGCCHSSDREKTKGKCEDDPFCVYKQEMSGHG 566
 QY 541 QCSGDCDCLDMDWTGYCNCTRTDTGCMSSNGLLSGRGKCEGSCVCTOPGSTGTCEK 600
 DB 567 QCSGDCDCLDMDWTGYCNCTRTDTGCMSSNGLLSGRGKCEGSCVCTOPGSTGTCEK 626
 QY 601 CPTCPDCTFKKEVECKKFPDRLALHDENCRNRCREISVYKELKTGDAVACTKNE 660
 DB 627 CPTCPDCTFKKEVECKKFPDRLALHDENCRNRCREISVYKELKTGDAVACTKNE 686
 QY 661 DDCVRFQYEDSSGKSLTYVEEPECCKGPDILVLSVGAITLLGLAALLTKLLIT 720
 DB 687 DDCVRFQYEDSSGKSLTYVEEPECCKGPDILVLSVGAITLLGLAALLTKLLIT 746
 QY 721 IHRKEFAKFEERARAKMDTANNPLKEATSTNTTNYRGT 762
 DB 747 IHRKEFAKFEERARAKMDTANNPLKEATSTNTTNYRGT 788

RESULT 4
 B36268
 platelet glycoprotein IIA-II - human (fragment)
 N:Alternate names: Integrin beta-3' chain
 C:Species: Homo sapiens (man)
 C:Date: 18-Jan-1991 #sequence_revision 13-Jan-1993 #text_change 20-Aug-1999
 C:Accession: B36268; A33907
 R:Zimlin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; P
 J. Biol. Chem. 265, 8590-8595, 1990
 A:Title: The genomic organization of platelet glycoprotein IIA.
 A:Reference number: A36268; MUID:90256778
 A:Accession: B36268
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-753 <ZIM>
 A:Cross-references: GB:J05427
 A:Note: the authors translated the codon GAT for residue 216 as Glu, GAT for residue
 R:van Kuppevelt, T.H.M.S.M.; Langui, L.R.; Galile, J.O.; Suzuki, S.; Ruoslahti, E.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5415-5418, 1989
 A:Title: An alternative cytoplasmic domain of the Integrin beta-3 subunit.
 A:Reference number: A33907; MUID:89315807
 A:Accession: A33907
 A:Molecule type: preliminary
 A:Status: preliminary
 A:Residues: 680-753 <VAN>
 A:Cross-references: GB:M25108; NID:g186502; PIDN:AAA36121.1; PID:g386833
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: cell adhesion; glycoprotein

Query Match 97.2% Score 4038; DB 2; Length 753;
 Best Local Similarity 99.6% Pred. No. 3, 3e-249;
 Matches 740; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNICITTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPPVSEA 61
 DB 1 PNICITTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEPPVSEA 60
 QY 62 RVLEDRPLSDKSGSSQVTOVSPORIALRLRPDDSKNFSIOVROYEYPPDIYLLML 121
 DB 61 RVLEDRPLSDKSGSSQVTOVSPORIALRLRPDDSKNFSIOVROYEYPPDIYLLML 120
 QY 122 YSMKDDLMSTIONLGRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENFCYDM 181
 DB 121 YSMKDDLMSTIONLGRKLTATQMRKLTSLRIGGAFVDRPVSPYMTISPEALENFCYDM 180
 QY 182 TTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGFDALMQATVDEKIGMNR 241
 DB 181 TTTCLPMFGYKHYLTLDQVTRNEEVKKOSVSRNDAPEGFDALMQATVDEKIGMNR 240
 QY 242 ASHLVFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLSQ 301
 DB 241 ASHLVFTTDAKTHIALDGRLAGIVQPNQGCHVSDNHYSASTTMDYPSLGLMTEKLSQ 300

```

QY 302 KNINLFAVTENVYLYQNVSELIPTGTGVLSDSSNVLLQIVDAVGKIRSKVELEVRD 361
DB 301 KNINLFAVTENVYLYQNVSELIPTGTGVLSDSSNVLLQIVDAVGKIRSKVELEVRD 360
QY 362 LPELISLSEFNATCLNNEVIFGLKSCGKLGITVVSFSIAKVRGCGQEKESFTIKPVGF 421
DB 361 LPELISLSEFNATCLNNEVIFGLKSCGKLGITVVSFSIAKVRGCGQEKESFTIKPVGF 420
QY 422 KDSLIYVTFDCCACAOAEPNSHRCNNGNGTFEGGVCRCGGMVLSQCESEEDYRPS 481
DB 421 KDSLIYVTFDCCACAOAEPNSHRCNNGNGTFEGGVCRCGGMVLSQCESEEDYRPS 480
QY 482 QODECSFREGQPYCSQSGECLCGQCVHSSDFGKITGKYCECDDEFCVARYKGMGSGHQ 541
DB 481 QODECSFREGQPYCSQSGECLCGQCVHSSDFGKITGKYCECDDEFCVARYKGMGSGHQ 540
QY 542 CSGGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRGKCEGSCVCIQPGSYGDTCEKC 601
DB 541 CSGGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRGKCEGSCVCIQPGSYGDTCEKC 600
QY 602 PTPDPACTFKKCEVECKKFPDRCALHDEMTNRCYCRDEISVYELKDTGKDAVNCYTKND 661
DB 601 PTPDPACTFKKCEVECKKFPDRCALHDEMTNRCYCRDEISVYELKDTGKDAVNCYTKND 660
QY 662 DCVAREQYEDSSGKSLVVEEPCPKGPDIIVLLSYMGAIIILIGLALLIKLLITI 721
DB 661 DCVAREQYEDSSGKSLVVEEPCPKGPDIIVLLSYMGAIIILIGLALLIKLLITI 720
QY 722 HDKKEFAKFEERAKAKMTANN 744
DB 721 HDKKEFAKFEERAKAKMTANN 743

```

RESULT 5

PN0509

Integrin beta-3 chain - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999

C:Accession: PN0509

R:Clontat, A.M.; Rosa, J.P.; Letourneur, F.; Poncez, M.; Rifat, S.

Biochem. Biophys. Res. Commun. 193, 771-778, 1993

A:Title: A comparative analysis of cDNA-derived sequences for rat and mouse beta-3 integ

A:Reference number: PN0509; MUID:93290675

A:Accession: PN0509

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-723 <CIR>

C:Superfamily: Integrin beta chain; laminin-type EGF-like homology

C:Keywords: cell adhesion; duplication; heterodimer; membrane protein

Query Match

Best Local Similarity 86.1%; Score 3577; DB 2; Length 723;

Matches 650; Conservative 42; Mismatches 31; Indels 0; Gaps 0;

```

QY 40 LKNLKNCAPESTIEFPVSEARVLEDRPLSDKSGSSQVTVSPQRIALRLRPDSSKN 99
DB 1 LKNLKNCAPESTIEFPVSEARVLEDRPLSDKSGSSQVTVSPQRIALRLRPDSSKN 99
QY 100 FSIQROVEDYVDIYLLMDLSYMKDLMSTONIGTKLATQMKLTNLRIGFGAFYDK 159
DB 61 FSIQROVEDYVDIYLLMDLSYMKDLMSTONIGTKLATQMKLTNLRIGFGAFYDK 120
QY 160 PVSPLYMISPPALENPPCYDMKTCLPMFGYKHVLTLDQVYRFNEEVKKOSVSNRPAP 219
DB 121 PVSPLYMISPPALENPPCYDMKTCLPMFGYKHVLTLDQVYRFNEEVKKOSVSNRPAP 180
QY 220 ECGFPAIMQATVCEKIGKRNDSHLVFTTDAKTHIALDGLAGIYVQNDGQCHVSGDN 279
DB 181 ECGFPAIMQATVCEKIGKRNDSHLVFTTDAKTHIALDGLAGIYVQNDGQCHVSGDN 240
QY 280 HYSASTMDYPSLGLMTEKLSOKNINLFAVTENVYLYQNVSELIPTGTGVLSDSSNV 339
DB 111 HYSASTMDYPSLGLMTEKLSOKNINLFAVTENVYLYQNVSELIPTGTGVLSDSSNV 339

```

```

DB 241 HYSASTMDYPSLGLMTEKLSOKNINLFAVTENVYLYQNVSELIPTGTGVLSDSSNV 300
QY 340 VQLIYDAVAGKIRSKVELEVRDLPBELSLSEFNATCLNNEVIFGLKSCGKLGITVVSFSI 399
DB 301 VQLIYDAVAGKIRSKVELEVRDLPBELSLSEFNATCLNNEVIFGLKSCGKLGITVVSFSI 360
QY 400 EAKVRGCGQEKESFTIKPVGFKDSLIYVTFDCCACAOAEPNSHRCNNGNGTFEGGVC 459
DB 361 EAKVRGCGQEKESFTIKPVGFKDSLIYVTFDCCACAOAEPNSHRCNNGNGTFEGGVC 420
QY 460 CRGGMVLSQCESEEDYRPSQODECSFREGQPYCSQSGECLCGQCVHSSDFGKITGK 519
DB 421 CRGGMVLSQCESEEDYRPSQODECSFREGQPYCSQSGECLCGQCVHSSDFGKITGK 480
QY 520 YECBDFSCVARKGKMGSGHQSCGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRG 579
DB 481 YECBDFSCVARKGKMGSGHQSCGDLCDSDMTGYCNCCTTRDTQMSNGLLCSRG 540
QY 580 KCEGSCVCIQPGSYGDTCEKCPTCPDPACTFKKCEVECKKFPDRCALHDEMTNRCYCRDEI 639
DB 541 KCEGSCVCIQPGSYGDTCEKCPTCPDPACTFKKCEVECKKFPDRCALHDEMTNRCYCRDEI 600
QY 640 ESKVLEKDTGKDAVNCYTKNDPVCVAREQYEDSSGKSLVVEEPCPKGPDIIVLLS 699
DB 601 ESKVLEKDTGKDAVNCYTKNDPVCVAREQYEDSSGKSLVVEEPCPKGPDIIVLLS 660
QY 700 VMGALLILGLALLIKLLITIHDKKEFAKFEERAKAKMTANNPLYKATSTFTNITY 759
DB 661 VMGALLILGLALLIKLLITIHDKKEFAKFEERAKAKMTANNPLYKATSTFTNITY 720
QY 760 RGT 762
DB 721 RGT 723

```

RESULT 6

S43534

Integrin beta3 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999

C:Accession: S43534; A49312

R:Minura, H.; Cao, X.; Ross, F.P.; Chiba, M.; Teitelbaum, S.L.

Endocrinology 134, 1061-1066, 1994

A:Title: 1,25-dihydroxyvitamin D(3) transcriptionally activates the beta(3)-integrin

A:Reference number: S43534; MUID:94164000

A:Accession: S43534

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <MIM>

A:Cross-references: EMBL:X72378; NID:9474038; PIDN:CAA51069.1; PID:9474039

A:Note: the authors translated the codon TTT for residue 174 as GLY, GGC for residue

R:Caio, X.; Ross, F.P.; Zhang, L.; Macdonald, P.N.; Chappel, J.; Teitelbaum, S.L.

J. Biol. Chem. 268, 27371-27380, 1993

A:Title: Cloning of the promoter for the avian integrin beta-3 subunit gene and its r

A:Reference number: A49312; MUID:94086557

A:Accession: A49312

A:Molecule type: DNA

A:Residues: 1-26 <CAO>

A:Cross-references: GB:X75348; NID:9441308; PIDN:CAA53095.1; PID:9452854

C:Superfamily: Integrin beta chain; laminin-type EGF-like homology

Query Match

Best Local Similarity 83.9%; Score 3484.5; DB 2; Length 781;

Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;

```

QY 1 GPNICTTGVSSCCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKNCAPESTIEFPVSE 60
DB 20 GSNICATRGVTSCKQCLAVSPMCAMCSAEVVAOSTPRCDLFAULLNGCGRPDIETPRSS 79
QY 61 ARVLEDRPLSDKSGSSQVTVSPQRIALRLRPDSSKNFSIYQVDEYVDIYLLMDL 120
DB 80 ITVLEDRPLSDKSGSSQVTVSPQRIALRLRPDSSKNFSIYQVDEYVDIYLLMDL 138

```


Oy	121	SYSMKDDLMASIONGDKTKATOMRKLTSLNRIGFGFVYKPPSPRYATISPPKLELNPICDM	180
Dd	139	SNSMKDDKLKIONLTKTAKLASEMRKLTSLNRIGFGAFVKKPISPPIYATISPPKAIKPCIEI	198
Oy	181	KTTCLPMGCKHVLTLDQVTRFNEEVRKQSVSRNRDAPEGGFDAIMQATYDEKIGMRN	240
Dd	199	GEKCLPMGCKHVLTLDQVTRFNEEVRKQSVSRNRDAPEGGFDAIMQATYDEKIGMRN	258
Oy	241	DASHLVTTTQAKTHIALDGRLAGIVQPNDOCHVGSQNMHSASTTMYPSLGLMTEKLS	300
Dd	259	DASHLVTTTQAKTHIALDGRLAGIVQPNDOCHIDKONFYSASTTLDYPSLGLMTEKLS	318
Oy	301	QKNILIFAVTENVNLYKONSSELIPGTTVGLSHDSSNNVLOLIYDAYGKIISKYELFVR	360
Dd	319	QKNILIFAVTDVTVGLVQNSSELIPGTTVGLSHDSSNNVLOLIYDAYGKIISKYELFVR	378
Oy	361	DLPELJLSFNAKTCLNNEYIPGLKCMGLKIGDIVSFSIEAKVRCCPOEKESFTIKFVG	420
Dd	379	DLPELJLSFNAKTCLNDEYITGLKCMGLKIGDIVSFSIEAKVRCCPOEBROKSPFIKFPVG	438
Oy	421	FKDSLIVQVTFPCDCACOAQAPNHSRGNNGTPEGVCVRCRGPWLSOCCESEEDRP	480
Dd	439	FKDSLIVVYNEDPCMSCESQAEANSSFCSSKNGSLIEGVCVRCNRPRLSHCECSEETNP	498
Oy	481	SQODECSPREGQPYVCSQREKCLGQCVCHSSDPFKITGKYCECDDFSCVRYKGEKCSGHG	540
Dd	499	SEQDNCSPPQPGPLCSQREKCEICGQCVCHGSDPGFVYTGKCECDDFSCVRYKGEKCSGHG	558
Oy	541	QCSGDCCLCDSMTGTGYCNCRTTRDTCHSSNGLLCSGRKCEKCSGCVCIQPSYDDETEK	600
Dd	559	QCSGDCCLCDSMTGTGYCNCRTTRDTCHSSNGLYCSGHGICGKCDICIQPSYGNTEK	618
Oy	601	CPTCPDACKFKKCEYCKKFKDGRGALHDEWTCNRYCRDREIESYKELKDPGKDAVNTKYNE	660
Dd	619	CPTCPDACKFKKCEYCKKFKDGRGALHDEWTCNRYCRDREIETQELGDKGDAVNTKYDE	678
Oy	661	DDCVYRFQYEDSSGKSLIYVEEPECPKGPDIYVLLSVMGAILLIGLAILLWKLIT	720
Dd	679	NDCVYRFQYEDSSGKSLIYVEEPECPKGPDIYVLLSVMGAILLIGLAILLWKLIT	738
Oy	721	IHDRKFAKFEERARAKWDTNANPLYKEATSTFNITTYRG	761
Dd	739	IHDRKFAKFEERARAKWDTNANPLYKEATSTFNITTYRG	779

Db ASQRKILTSNLRIGFGAIVFDVKPSPYMIIISPQAIKNPCYNNMKNAACLPMIFGCKHVLITTD 120
Qy 199 QVTRPNEEVKKQOSYSRNNDAPEGGFDAIMQATVCDEKXIGANDSLLLVFTTDDAKTHIAL 258
Db 121 QVSFRNEVEKKQOSYSRRNDAPEGGFDAIMQATVCDEKXIGANDSHLLVFETDPAKTHIAL 180
Qy 259 DGRLAGIYQPNDGGCHGSODNHYSASTTMDYPSLGLMTERTKSQNLNLFVATSNVNVLY 318
Db 181 DGRLAGIYLPNDGGCHIGTDNMHYSASTTMDYPSLGLMTERTKSQNLNLFVATSNVNSLY 240
Qy 319 QNYSSELIPGTTVGVLSMDSSNVQLIYDAYGKIRSKVELFVRDLPEELSLSFNATCLNNE 378
Db 241 QNYSSELIPGTTVGVLSDDSSNVQLIYDAYGKIRSKVELFVRDLPEELSLSFNATCLNNE 300
Qy 379 VIPGLKSGMGKIKIDITYSFSIEAVYRGCPOKEKESFTTKPKPGFSDSLIVQTFECCDACQ 438
Db 301 VIPGLKSCVGLIKIDITYSFSIEAVYRGCPOKEKESFTTKPKPGFSDSLIVQTFECCDACQ 360
Qy 439 AQAEBSNRHCNNNGNGTEFCGCGVRCGPWLISQCCESEEDYPPSOODECSPREGQPVCSQR 498
Db 361 AFAQDPSPRCNNNGNGTEFCGCGVRCRDQGLSMCMCSEEDYPPSOODECSPREGQPICSOR 420
Qy 499 GECLCGQCVCCHSSDFGKITGKYRCECDPFSCYRVYGEWCSGHGCCSCGDCLCDSPWTGYTC 558
Db 421 GECLCGQCVCCHSSDFGKITGKYRCECDPFSCYRVYGEWCSGHGCCSCGDCVDSPTWTGYTC 480
Qy 559 NCTRTDTVMCSNGLICSGRGKCGCGCVCIOPSYGTCCKCPTPCAPCFEFKKECYECK 618
Db 481 NCTRTDTVMCSNGLICSGRGKCGCGCVCIOPSYGTCCKCPTPCAPCFEFKKECYECK 540
Qy 619 KEDRGALHDEMTCNNRYCDEDEIESYKELKDTPKDKAVNCTYYKNKEDDVAFYQYEDPSGKSI 678
Db 541 KFNQGTLHEEMTCNNRYCDEDEIESYKELKDTPKDKAVNCTYYKNKEDDVAFYQYEDTSGNAV 600
Qy 679 LYVVEEPCKPGPDILLVLLSVMGAILLGALAILMLKLTTIHDRKEFAKFEEBRARAK 738
Db 601 LYVVEEPCKPGPDILLVLLSVMGAILLGALAILMLKLTTIHDRKEFAKFEEBRARAK 660
Qy 739 WDTANNPLXYKATSTFTNIT 758
Db 661 WDTANNPLXYKATSTFTNIT 680

RESULT 8
151530
Integrin beta-3 subunit - African clawed frog
c:Species: Xenopus laevis (African clawed frog)
c:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Aug-1999
c:Accession: I51530
R:Ransom, D.G.; Hens, M.D.; Deslmonne, D.W.
Dev. Biol. 160, 265-275, 1993
A>Title: Integrin expression in early amphibian embryos: cDNA cloning and characterization
A:Reference number: I51530; MUID:94040374
A:Accession: I51530
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-788 <RAN>
c:Cross-references: GB:U3591; NID:g467811; PID:NAA17427.J; PID:g467812
c:Superfamily: Integrin beta chain; laminin-type EGF-like homology

Db 738 LLAIMMLVTHIDREFAFQSESRNARTEMANPLRKRPISHTHYDFTFNKNSYNGT 797

RESULT 10
A37057
Integrin beta-6 chain - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 19-Jan-2001
C:Accession: A37057; 169201
R:Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
J. Biol. Chem. 265, 11502-11507, 1990
A:Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) identified
A:Reference number: A37057; MUID:90307659
A:Accession: A37057
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-788 <SHE>
A:Cross-references: GB:M35198; GB:J05522; NID:g186506; PID:AAA3612.1; PID:g186507
R:Jiang, W.M.; Jenkins, D.; Yuan, Q.; Leung, E.; Choo, K.H.; Watson, J.D.; Kristensen, G
Int. Immunol. 4, 1031-1040, 1992
A:Title: The gene organization of the human beta 7 subunit, the common beta subunit of t
A:Reference number: 154749; MUID:93002753
A:Accession: 169201
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 116-157, 'R', 159-197 <JIA>
A:Cross-references: GB:545380; NID:9257588; PID:AA823690.1; PID:9257589
C:Genetics:
A:Gene: GDB:ITGB6
A:Cross-references: GDB:131392; OMIM:147558
A:Map position: 2pter-qter
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: blocked amino end; cell adhesion; cytoskeleton; glycoprotein; lipoprotein; p
F:708-730/Domain: transmembrane #status predicted <RMB>
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:7/Binding site: palmitate (Cys) (covalent) #status predicted
F:16,48,97,260,387,396,463,471/Binding site: carboxyrate (Asn) (covalent) #status predi

Query Match 51.9%; Score 2156; DB 2; Length 788;
Best Local Similarity 50.6%; Pred. No. 1.9e-129;
Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

Qy 1 GPNICITR-----GVSSCOQCLAVSPWCAMCSDAL--PLG-SPRCDLKENLKNCADE 52
Db 14 GRNDSRTMLCLGAGTCECDLIGPOCAMCAQENFTHPSGVECDTPANLLACQLN 73
Qy 53 STEEPSEARVLEDRPLSKSGSSQVQVSPQIALRLPRDDSKNFSIQVQVQVEDPY 112
Db 74 FLENPSOVELKMKPLSVGRKNSSDIYQIAQSLILKLRGGAQTLOVHVRQTEDEPV 133
Qy 113 DIYVLDLSYMKDLSIQLNLTQKLTATQMRKLTSLRIGFAGFVQKPYSPMYTSPPEA 172
Db 134 DLYVLDLSAMDDLTNTIKELSGSKSKMSKLTNFRIGFSGFVQKPYSPMYTSPPEA 192
Qy 173 LBNPCDMMKTTCLPMFGYHVLTLDOVTRFNEEYKQKOSVRRNDAPESGFDALMQATVC 232
Db 193 IANPCSSIPYFCILPFGFHIPLTLIDAEFNEIYKQKISANIDTPREGGFALMQAAVC 252
Qy 223 DEKIGRNDASHLLVTTDAKTHIALDGLAGIYQVNDQCVGSPNHSASTMDYPSL 292
Db 253 KETIGRNDASHLLVTTDAKTHIALDGLAGIYQVNDQCVGSPNHSASTMDYPSL 312
Qy 293 GLMTEKLSOKNINILFVAVTEVNVNLQYNSLIPGTIVYVLSMDSNNVQLLIVDAVKIR 352
Db 313 GQIDRLVQNNVLLFAVYQEQVHLKLENAKILPGATVGLLOKQSNINILQILIAVEELR 372
Qy 353 SKVELVRLPEELSLFNAATCLNNEVIRGLKSCMKLIGTVVSIEAKVGGCQOEK 412
Db 373 SEVELVRLPEELSLFNAATCLNNEVIRGLKSCMKLIGTVVSIEAKVGGCQOEK 431
Qy 413 SFIRKVGKDSLIVQVTEPCDCAQOAEPSNHRNNGNGTFEGGVCGGGMVGSQOE 472

Db 432 HIITKPVGLGALBELVSPENCOCQKEVEVNSSKCHHGNSFEGCVCACHPGMRCE 491

Qy 473 CSEEDYRPSQODECSPRGQPVCSQGRGELCGQCVCHSDPGKITGKYCECDDESVYRK 532
Db 492 CGED---MLSTDSCKEAPDRHPSGSGRGDCYGGCICHLSPYGNLGYPCOCDDNFSYRH 548
Qy 533 GEMCSGHGQSCGCDLSDMTGYCNCYTRTPDMSSNGLGSGRGKCEGSCVCIOPG 592
Db 549 GLLCGNGDCDCCGRCVCSGWTGECYCNCTSTDCVSDGDLGSGRDCVCGKCVCTNP 608
Qy 593 SYGDCEKCPKCPKPCPKKCEVCKKPRDGLADEMTCNRYCDELTESYELKEDTKD 651
Db 609 ASGPCEKCPKCPKPCPKKCEVCKKPRDGLADEMTCNRYCDELTESYELKEDTKD 666
Qy 652 AVNCTYKNECDVAVFOYEDSSGKSLVYVEEPCPKGPDILVLLSYMGAILLGLAA 711
Db 667 SVSCGLGECNCLITFLTITDNEKTIHSHNEKDCPRPPIPMILGLVSLATLILVIL 726
Qy 712 LLIKMLTTHIDREFAFQSESRNARTEMANPLRKRPISHTHYDFTFNKNSYNGT 760
Db 727 LCIWLLVSPFDRKREVAFAERAKAKWQGTNPILYKSTPFRNVYK 775

RESULT 11
J02005
Integrin beta-5 chain - baboon
C:Species: Papio sp. (baboon)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 21-Jan-2000
C:Accession: J02005
R:Shoji, M.; Hayzer, D.J.; Kim, T.M.; Runge, M.S.; Hanson, S.R.
Gene 133, 307-308, 1993
A:Title: Human and baboon Integrin beta 5 subunit-encoding mRNAs have alternative pol
A:Reference number: J02005; MUID:94040831
A:Accession: J02005
A:Molecule type: mRNA
A:Residues: 1-656 <SHO>
A:Note: The authors translated the codon AGA for residue 454 as Lys, CNA for residue
C:Comment: This protein is a predominant subunit for the vitronectin receptor in babo
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: cell adhesion
F:320-370/Domain: laminin-type EGF-like homology <LEG>

Query Match 48.0%; Score 1994.5; DB 2; Length 656;
Best Local Similarity 55.6%; Pred. No. 2.9e-119;
Matches 365; Conservative 102; Mismatches 175; Indels 15; Gaps 7;

Qy 118 MDLSYMKDDLSIQLNLTQKLTATQMRKLTSLRIGFAGFVQKPYSPMYTSPPEALNPC 177
Db 1 MDLSYMKDDLSIQLNLTQKLTATQMRKLTSLRIGFAGFVQKPYSPMYTSPPEALNPC 177
Qy 178 --YDKRTTCLPMFGYHVLTLDOVTRFNEEYKQKOSVRRNDAPESGFDALMQATVCDEK 235
Db 60 IGYKLFPCVCSFGRHILPLTDVDSNEEYKQKOSVRRNDAPESGFDALMQATVCDEK 235
Qy 236 IGRNDASHLLVTTDAKTHIALDGLAGIYQVNDQCVGSPNHSASTMDYPSLGLM 295
Db 120 IGRNDASHLLVTTDAKTHIALDGLAGIYQVNDQCVGSPNHSASTMDYPSLGLM 295
Qy 296 TEKLSOKNINILFVAVTEVNVNLQYNSLIPGTIVYVLSMDSNNVQLLIVDAVKIR 352
Db 180 GELKLAENNINILFVAVTKNHYMLYKFNFTALIPGTVEILIDGSKNIIQILINAVYSISKV 352
Qy 356 ELEVDRLPEELSLFNAATCLNNEVIRGLKSCMKLIGTVVSIEAKVGGCQOEK 414
Db 240 ELVWDQDEDLNFTATACQGVSYPRGQKCEGLKIDTASFEVSAVRCSRSHTEHVF 299
Qy 415 TIRVGFKDSLIVQVTEPCDCAQOAEPSNHRNNGNGTFEGGVCGGGMVGSQOE 474
Db 300 ALOPGCHDSLELVGVTNCTCGCGSVGLEPNSARC--SGTGYVGLCLCEPSVGLTTRCEQ 358
Qy 475 BEDYRPSQODECSPRGQPVCSQGRGELCGQCVCHSDPGKITGKYCECDDESVYRK 534

Db 359 DGENH5YQNLCDRTBEGKPLCSGRGDCSCNOCSEFSEFGKIVGPCEDCNFSCARKGV 418
 QY 535 MCSGHGCGCGDCLCDSDMTGYCNCCTTRDTGMSNGLLCSGRGCEGSCVCIOPG5V 594
 Db 419 LCSGHGCHGCECKCHAGYIGDNCNCSTLSTDRGRDGOICSRGHCLOCOCCTERPA 478
 QY 595 GDTCEKCPDACEKCECKEVECKKFDRCALHDENTCNRYCRDEIES-VKELDKTGDAV 653
 Db 479 GEMCEKCPDACEKCEKCECKEVECKKFDRCALHDENTCNRYCRDEIES-VKELDKTGDAV 537
 QY 654 NCYKKNDDCVNRFQYEDSSGSLLYVEPEPCPKPDILVLLVSMGALLIGLALL 713
 Db 538 LCFETAKDCVMMFTYELPSGKSNLTLYLREPECGMTNPMATLLVAGSLLVGLALLA 597
 QY 714 IMKLLITHDRKEFAKEEERAKMDTANNPLYKEATSTFT-----NITRG 762
 Db 598 IMKLLVTHDRREPAKQSRERARERASNPILYKRPISHTYDTFENKKNKSYNGT 654

RESULT 12

Integrin, band 3 precursor - chicken
 N:Alternate names: CSAR antigen; JG22 antigen; RGD-receptor
 C:Species: Gallus gallus (chicken)
 C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 23-Jul-1999
 C:Accession: A23947
 R:Rankun, J.W.; Desimone, D.W.; Fonda, D.; Patel, R.S.; Buck, C.; Horwitz, A.F.; Hynes, Cell 46, 271-282, 1986
 A:Title: Structure of integrin, a glycoprotein involved in the transmembrane linkage bet
 A:Reference number: A23947; MUID:86245073
 A:Accession: A23947
 A:Molecule type: mRNA
 A:Residues: 1-803 <TAM>
 A:Cross-references: GB:M14049; NID:9212213; PIDN:AAA48926.1; PID:9212214
 A:Experimental source: embryonic fibroblasts
 A:Note: the amino end of the mature protein is blocked
 C:Comment: Integrin, an integral plasma membrane complex of three (two in mammalian cell
 C:Comment: This transmembrane complex may be the target of oncogenic transformation that
 C:Comment: The cytoplasmic domain of this subunit contains a potential tyrosine-kinase pho
 C:Comment: The extracellular domain of this protein, like many membrane receptors, conta
 C:Keywords: Cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-803/Product: Integrin, band 3 #status predicted <MAM>
 F:25-733/Domain: extracellular #status predicted <EXT>
 F:467-654/Region: cysteine-rich
 F:482-564,565-650/Region: duplication
 F:734-756/Domain: transmembrane #status predicted <MEM>
 F:757-803/Domain: intracellular #status predicted <INT>
 F:216,273,367,410,421,433,445,486,525,589,624,674/Binding site: carbohydrate (Asn) (cova
 F:788/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 44.5%; Score 1846.5; DB 1: Length 803;
 Best Local Similarity 45.1%; Pred. No. 9,4e-110;
 Matches 353; Conservative 136; Mismatches 266; Indels 27; Gaps 15;

QY 1 GPNICTRGVSSCOCCIAVSPMCAMC-SDEALPLGSP---RCOLKENLLKDNCAPESTIEF 56
 Db 27 GGSDCIRANAKKSCGECIOAGPNCGWCKKIDFLOEGEPTAKRCDLALAKSKGCEPDIEDN 86
 QY 57 PVSEARVLEDRPLSDKSGDSSO-----VTQVSPRIALRLRPDDSKNFSIQVRQVEDPY 111
 Db 87 PRGSKRVLEDRVETNRKIGAAEKRLPEALIQIPQKLVLIQLRGEGPQTFSLMKRRADXP 146
 QY 112 VDIYLLDLSYSKMDLMSIQNTIGTLATOMKRLTSNLRIGAGVADKVPSPYMTSPPE 171
 Db 147 IDLYLLDLSYSKMDLENKVSIGTALMREMEKITSDFRIGSFSEKTYMPTISTTPAK 206
 QY 172 ALNPPCDMKTGCLPMGEGYHVLTLDDVTRNEEYKOSYSNRRAPGSPAIQAV 231
 Db 207 LRRPCTG-DONCTSPSPSYKNVLSITSEGNKFNELVKGKHISGNLSPGSGDALIMQAV 264

QY 232 CDKTIQWRNDASHLLVFTTDAKTHIALDRLAGIVQPDGQCHGSDNHNYSATTTMDYPS 291
 Db 265 CDDQJGWRN-VYRLLEVDAGFHEAGDKLGIGIPLVDGKCHL-ENNMTM5HYDDPS 322
 QY 292 LGLMTEKLSQKINILIFAVTEVNVVLQYNSLITGTVGLSDSSVAVLIQIYDAKGI 351
 Db 323 IAHLVQKLSENNIGTIFAVTEEVAVYKELNKLPSKAVGLTSSNSVAVLIQIYDAKGI 382
 QY 352 RKRVLEVDLPEEELSLSPNATCLN-NEVIFGLKSCGMLKIPVFSIARVACPOE 409
 Db 383 SBEVILENSKLPKEVITISKYSCKNGVNDTQEDGKCNISIGDEVREIVTANNECKK 442
 QY 410 -KEKSFITKPVGKDSLIVQVTFDCCACQQAEPNSHRCNNGNGTEFCGVCRCGPMWG 468
 Db 443 GNETIKITPLGFTBEVEIHIOFIDCLQSEGEPSBACHDGNTEFCGACRCNEGRIG 502
 QY 469 SOCCESEEDYRSQODE-CSPREQPVCSOGEICLCGCGCVCHSSPFGK--ITGVCCEDD 525
 Db 503 RLCESTIDEVNSEDMDAYCIRENSTEICSNNGEICLGCGCVCKKRNTNEYVSGKICECN 562
 QY 526 FSCVRYKGMCSGHGQSCGDCIDSDMTGYCNCCTTRDTGMSNGLLCSGRGCEGCS 585
 Db 563 FNCDSRNGLLCGNGICRCVCEPNTFGSACDCLDTPCOMANGQICNGRGCEGT 622
 QY 586 CYCIOPG5VGTCEKCPDACEKCECKEVECKKFDRCALHDENTCNRYC----RDEIES 641
 Db 623 CNCTDPEKFGPCEGCGTCLGVCARHKDCVOCRAFEK--EKETCSQCEMHFMNTRYES 680
 QY 642 VKEU-KDTCGKDAV-NCYTKNEDDCVVRFOYEDSSGKSLLYVEPEPCPKPDILVLLS 699
 Db 681 RGLKQPNHPDPLSHCKEKDVGDCMFTYISVNSNGEASVHVEPEPCSPDIIPIYAG 740
 QY 700 VNGAILLIGLALLIMKLLITHDRKEFAKEEERAKMDTANNPLYKEATSTFTNITV 759
 Db 741 VVAGIVLIGLALLIMKLLITHDRREPAKQSRERARERASNPILYKRPISHTYDTFENKKNKSYNGT 800
 QY 760 RG 761
 Db 801 EG 802

RESULT 13

fibronectin receptor beta chain precursor - human
 N:Alternate names: CD29 antigen; Integrin beta-1 chain
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999
 C:Accession: B27079
 R:Argaves, W.S.; Suzuki, S.; Arai, H.; Thompson, K.; Pierschbacher, M.D.; Ruoslahti, J. Cell Biol. 105, 1183-1190, 1987
 A:Title: Amino acid sequence of the human fibronectin receptor.
 A:Reference number: A27079; MUID:88007843
 A:Accession: B27079
 A:Molecule type: mRNA
 A:Residues: 1-798 <ARC>
 A:Cross-references: EMBL:X07979; NID:931441; PIDN:CAA30790.1; PID:931442
 C:Genetics:
 A:Gene: GDB:ITGB1; FNRB
 A:Cross-references: GDB:118732; OMIM:135630
 A:Map position: 10p11.2-10p11.2
 C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C:Keywords: alternative splicing; duplication; heterodimer; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-798/Product: fibronectin receptor beta chain #status predicted <MAM>

Query Match 44.1%; Score 1831; DB 2: Length 798;
 Best Local Similarity 44.9%; Pred. No. 9e-109;
 Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

QY 3 NICTTRGVSSCOCCIAVSPMCAMCSD-----EALPLGSPRDLKENLLKDNCAPESTIEF 57
 Db 25 NRCIRANAKKSCGECIOAGPNCGWCTNFTLOEGEMT-SARCDLDELAKKKGCPDIEDN 83

F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-798/Product: integrin beta-1 chain #status predicted <MAT>

Query Match 43.8%; Score 1821; DB 2; Length 798;
Best Local Similarity 44.4%; Pred. No. 3.9e-108;
Matches 346; Conservative 137; Mismatches 269; Indels 28; Gaps 13;

```
QY 3 NICTRGVSSCQCLAVSPKANCSD----EALPLGSPRCDLKENILKDNCAPESTTEPP 57
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 NRCLKANAKSCGECIAGPNCGWCTNTTFLQEGMPT-SARCDLLEALKKKGGCPSDIENP 83
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 VSEARVLEDRPLSDKSGDSQ-----VTQVSPQIALRLRPDSSKNFSTQVRQVEDYDV 112
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 RGSQTIKKKNKNTVNRKSKMAEKLRPEDITQIQPQLLKLRSQEPQKFTLKKFRADYPI 143
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 113 DIYYLMDLSYMKDLSIONLSTKLATQMRKLTSLNRIGAFVDPKPSYPMYISPPPEA 172
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 144 DLYYLMDSLYSMKDLENVKSLGTDLMNEMRRTSDFRIGFSFVKTYMPTISTTPAK- 202
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 LEMPCYDMKTTCLPMEGYKHVLTLDQVTRFNEEVKQSVSRNRDAPEGFPDAIMQATVC 232
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 LRNPC-TSQNCTSPESYNNVSLTDREFFNELVGOQRISGNDLSPGEGFDIMQVAVC 261
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 DEKIGRNDASHLLVETTPAKTHIALDRGLAQVQNDGQCHVGSNDNHSASTMDYPSL 292
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 262 GSLIGRN-VTRLVTFSTAGHFGADGKGLGIVLPNDQCHL-ENNYTMSHYDYDPSI 319
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 293 GLMTEKLSOKNINLFAVTENVNLYQNSSELPGTWGLSMDSSNYLQILVDAYGKIR 352
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 320 AHLVOKLSENNIQTIFAATEEFQPYKELKNLIPKSAVGTLSGNSNVIQILIIDAYNSLS 379
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 353 SKVELVRLPRELSLSFNATCLN--NEYIPLKSCMGKIGDIVSFSIEAKVRGCPQEK 410
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 380 SEVILNSKLPDGVITINKSYCKNGVNGTGENGRKCSNISIGDEVOFEISTANKCPNKE 439
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 EKSFYIKPVGFDSLIYOVTFDDCAQAOAEPNSHRGNNGNGTPECGVCRGPGWLGSO 470
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 440 SETIKIKPLGTFEEVEVYVLOFLCKNCOSHGIPASPKCHEGNGTFEGCAGCRNCGRYGRH 499
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 CECSEEDYRPSQODE-CSPREGQPYVCSQRGBCLGQCVCHSSDFGK--ITGKYCECDDES 527
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 500 CECSTDEVNSEDMDAYCKRKNSSSEICSNNGECVCGQCVCRRKRDNTNIEYSGKFCEDNFN 559
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 528 CVRYGEMSGHGOCSGDCLDSDMTGYCNCITTRDTCMSSNGLLCGRKCCEGSCV 587
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 560 CDRSNGLICGNGVCRCVCEYCPYVTSACDCSLDTGPCLASNGOICNGRIGCEGACK 619
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 588 CIQPSYGDTCCKPTGPDACCTFKKCEVCCKFDRGALHDENTCNRYCR---DEISVK-643
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 620 CTDPKFGPTCETCOTGLGVAEHKECVQCAFNKGEKD--TCAQECSHENLTKVESRE 677
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 644 ELKDTGK--DAVNCTYNEDCVRFQYEDSSGKSLIYVEEPECPKGPDLIVLLSV 701
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 678 KLPQPVQDVPYTHCKEKDIDCMWFYFTYSVAGNNEAIVHVEPDCPTGPDIPIVAGV 737
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 702 GAILLIGLALLIKLITTHDRKEFAKFEERARAKMDTANNPLYEATSTFTNITYRG 761
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 738 AGIYLIGLALLIKLIMITHDRREFAKFEKEMAKMDTGENDPIYKSAVTIVANPKYEG 797
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 19, 2002, 12:24:02
Job time: 9709 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 10:39:46 ; Search time 40.44 Seconds

(without alignments)
729,582 Million cell updates/sec

Title: US-09-673-302a-1

Perfect score: 4154
Sequence: 1 GPNICTTNGVSSCQCLAVS.....NNPLYKATSTFTNITYRGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4119	99.2	788	ITB3_HUMAN	P05106 homo sapien
2	3839	92.4	787	ITB3_MOUSE	O54890 mus musculu
3	2343	56.4	798	ITB5_MOUSE	O70309 mus musculu
4	2339	56.3	799	ITB5_HUMAN	P18084 homo sapien
5	2177	52.4	787	ITB6_MOUSE	O90019 mus musculu
6	2170.5	52.3	788	ITB6_HUMAN	P18564 homo sapien
7	1989.5	47.9	655	ITB5_PAPCT	O07441 papio cynoc
8	1846.5	44.5	803	ITB1_CHICK	P07228 gallus gall
9	1833	44.1	798	ITB1_FELCA	P53713 felis silve
10	1831	44.1	798	ITB1_HUMAN	P05556 homo sapien
11	1826.5	44.0	798	ITB0_MOUSE	P12607 xenopus lae
12	1821	43.8	798	ITB1_MOUSE	P16005 mus musculu
13	1818.5	43.8	798	ITB1_XENLA	P16006 xenopus lae
14	1817	43.7	773	ITB1_BOVIN	P53712 bos taurus
15	1801.5	43.4	799	ITB1_RAT	P49134 rattus norv
16	1676	40.3	577	ITB6_CAVPO	P18563 cavia porce
17	1579	38.0	809	PAT3_CAEEL	O27874 caenorhadi
18	1537	37.0	845	ITB2_DROME	P11584 drosophila
19	1520.5	36.6	769	ITB2_BOVIN	P33592 bos taurus
20	1515	36.5	771	ITB2_MOUSE	P11835 mus musculu
21	1513.5	36.4	769	ITB2_HUMAN	P05107 homo sapien
22	1509.5	36.3	769	ITB2_PIG	P53714 sus scrofa
23	1506	36.3	806	ITB7_MOUSE	P26011 mus musculu
24	1502.5	36.2	798	ITB7_HUMAN	P26010 homo sapien
25	1331	33.0	769	ITB8_RABIT	P26013 oryctolagus
26	1326.5	31.9	768	ITB8_HUMAN	P26012 homo sapien
27	1184	28.5	1822	ITB4_HUMAN	P16144 homo sapien
28	1154.5	26.8	1807	ITB4_RAT	O64633 rattus norv
29	1097	26.4	799	ITB1_DROME	O27591 drosophila
30	260.5	6.3	4289	TENX_HUMAN	P22105 homo sapien
31	259	6.2	2201	TENX_HUMAN	P24821 homo sapien
32	255	6.1	1746	TENX_PIG	O29116 sus scrofa
33	252.5	6.1	1808	TENX_CHICK	P10039 gallus gall

ALIGNMENTS

RESULT	ID	STANDARD	PRT	788 AA
1	ITB3_HUMAN			
34	239.5	5.8	830	SREC_HUMAN
35	214	5.2	3106	LMH2_MOUSE
36	212	5.1	3672	LMH2_CAEEL
37	211	5.1	722	DLH1_MOUSE
38	210	5.1	714	DLH1_RAT
39	207	5.0	2437	NOTC_BRARE
40	205	4.9	2871	FBN1_MOUSE
41	204	4.9	723	DLH1_HUMAN
42	202	4.9	2531	NTCL_MOUSE
43	201.5	4.9	2139	NTCL_MOUSE
44	201	4.8	2871	FBN1_PIG
45	200.5	4.8	2444	NTCL_HUMAN
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RM	SEQUENCE FROM N.A. (ISOFORM BETA-3A).			
RP	MEDLINE=87165991; PubMed=3494014.			
RX	Fitzgerald L.A., Steiner B., Rall S.C. Jr., Lo S., Phillips D.R.,			
RA	"Protein sequence of endothelial glycoprotein Ila derived from a			
RT	cDNA clone, identity with platelet glycoprotein Ila and similarly			
RT	to "Integrin".			
RL	J. Biol. Chem. 262:3936-3939(1987).			
RM	SEQUENCE FROM N.A. (ISOFORM BETA-3A).			
RP	MEDLINE=90265363; PubMed=2345548.			
RX	Frerich P., Uzan G., Thevenon D., Denarter E., Prandini M.H.,			
RA	Marguerie G.,			
RT	"GPIIb and GPIIla amino acid sequences deduced from human			
RT	megakaryocyte cDNAs."			
RL	Mol. Biol. Rep. 14:27-33(1990).			
RM	SEQUENCE FROM N.A. (ISOFORM BETA-3A).			
RP	MEDLINE=88213696; PubMed=2452834.			
RX	Zimrin A.B., Elisman R., Villalre G., Schwartz E., Bennett J.S.,			
RA	Poncz M.,			
RT	"Structure of platelet glycoprotein Ila. A common subunit for two			
RT	different membrane receptors."			
RL	J. Clin. Invest. 81:1470-1475(1988).			
RM	SEQUENCE FROM N.A. (ISOFORM BETA-3A).			
RP	MEDLINE=90256778; PubMed=2341395.			
RX	Zimrin A.B., Gidwitz S., Lord S., Schwartz E., Bennett J.S.,			
RA	White G.C. II, Poncz M.,			
RT	"The genomic organization of platelet glycoprotein Ila."			
RL	J. Biol. Chem. 265:8590-8595(1990).			
RM	SEQUENCE FROM N.A. (ISOFORM BETA-3C).			
RP	TISSUE=Osteoclastoma;			
RX	MEDLINE=97341177; PubMed=9195946;			
RA	Kumar C.S., James I.E., Wong A., Kwang V., Feld J.A.,			
RT	Nathaniaganti P., Connor J.R., Eichman C., Ali F., Hwang S.M.,			
RT	Rieman D.J., Drake F.H., Gowen M.,			
RL	"Cloning and characterization of a novel integrin beta3 subunit."			
RL	J. Biol. Chem. 272:16390-16397(1997).			
RP	SEQUENCE OF 1-26 FROM N.A.			

- RC TISSUE-Blood;
RX MEDLINE=94129007; PubMed=8298129;
RA Villa-Garcia M., Li L., Riely G., Bray P.F.;
RT "Isolation and characterization of a TATA-less promoter for the human
RT beta 3 integrin gene.";
RL Blood 83:668-676(1994).
- RN SEQUENCE OF 27-788 FROM N.A. (ISOFORM BETA-3A).
RX MEDLINE=91009291; PubMed=2145280;
RA Lanza F., Kieffer N., Phillips D.R., Fitzgerald L.A.;
RT "Characterization of the human platelet glycoprotein IIa gene.
RT Comparison with the fibronectin receptor beta-subunit gene.";
RL J. Biol. Chem. 265:18098-18103(1990).
- RN [8]
RX SEQUENCE OF 122-204 FROM N.A.
RX MEDLINE=93002753; PubMed=1382574;
RA Jiang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
RA Kristiansen G.W.;
RT "The gene organization of the human beta 7 subunit, the common beta
RT subunit of the leukocyte integrins HML-1 and LPM-1.";
RL Int. Immunol. 4:1031-1040(1992).
- RN [9]
RX PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-3B).
RX TISSUE-Placenta;
RX MEDLINE=89315807; PubMed=2787511;
RA Van Kuppevelt T.H.M.S.M., Langui L.R., Gallit J.O., Suzuki S.,
RA Ruostalahti E.;
RT "An alternative cytoplasmic domain of the integrin beta 3 subunit.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5415-5418(1989).
- RN [10]
RX SEQUENCE OF 218-234 AND 439-443.
RX MEDLINE=87101510; PubMed=3801670;
RA Hiraiwa A., Matsukage A., Shiku H., Takahashi T., Naito K., Yamada K.;
RT "Purification and partial amino acid sequence of human platelet
RT membrane glycoproteins IIb and IIa.";
RL Blood 69:560-564(1987).
- RN [11]
RX PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=91158732; PubMed=2001252;
RA Calvete J.J., Henschen A., Gonzalez-Rodriguez J.;
RT "Assignment of disulphide bonds in human platelet GPIIb. A
RT disulphide pattern for the beta-subunits of the integrin family.";
RL Biochem. J. 274:63-71(1991).
- RN [12]
RX PHOSPHORYLATION AT TYR-773 AND TYR-785 (ISOFORM BETA-3A).
RX MEDLINE=96210016; PubMed=8631894;
RA Law D.A., Nannizzi-Alaimo L., Phillips D.R.;
RT "Outside-in integrin signal transduction. Alpha IIb beta 3-(GP IIb
RT IIb) tyrosine phosphorylation induced by platelet aggregation.";
RL J. Biol. Chem. 271:10811-10815(1996).
- RN [13]
RX X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 27-718.
RX MEDLINE=21482770; PubMed=11546839;
RA Xiong J.P., Stehle T., Dieffenbach B., Zhang R., Dunker R., Scott D.L.,
RA Joachimiak A., Goodman S.L., Arnout M.A.;
RT "Crystal structure of the extracellular segment of integrin alpha
RT Vbeta3.";
RL Science 294:339-345(2001).
- RN [14]
RX VARIANT HPA-1 (PL(A)).
RX MEDLINE=89214713; PubMed=2565345;
RA Newman P.J., Derbes R.S., Aster R.H.;
RT "The human platelet alloantigens, PLAI and PLAI2, are associated with
RT glycoprotein IIa, and are distinguishable by DNA typing.";
RL J. Clin. Invest. 83:1778-1781(1989).
- RN [15]
RX VARIANT HPA-4 (PEN).
RX MEDLINE=93055444; PubMed=1430225;
RA Wang R., Furuhata K., McFarland J.G., Friedman K., Aster R.H.,
RA Newman P.J.;
RT "An amino acid polymorphism within the RGD binding domain of platelet
RT membrane glycoprotein IIa is responsible for the formation of the
- RT Pena/Pemb alloantigen system.";
RL J. Clin. Invest. 90:2038-2043(1992).
- RN [16]
RX VARIANT MO-1.
RX MEDLINE=93112977; PubMed=8093349;
RA Kuljpers R.W.A.M., Simsek S., Faber N.M., Goldschmeding R.,
RA van Weerkerken R.K.V., von Dem Borne A.E.G.K.;
RT "Single point mutation in human glycoprotein IIa is associated with
RT a new platelet-specific alloantigen (Mo) involved in neonatal
RT alloimmune thrombocytopenia.";
RL Blood 81:70-76(1993).
- RN [17]
RX VARIANT CA/TU.
RX MEDLINE=94060373; PubMed=7694683;
RA Wang R., McFarland J.G., Kekomaki R., Newman P.J.;
RT "Amino acid 489 is encoded by a mutational 'hot spot' on the beta 3
RT integrin chain: the CA/TU human platelet alloantigen system.";
RL Blood 82:3386-3391(1993).
- RN [18]
RX VARIANT SR(A).
RX MEDLINE=94179229; PubMed=8132570;
RA Santoso S., Kalb R., Kroll H., Walika M., Kiefel V.,
RA Mueller-Eckhardt C., Newman P.J.;
RT "A point mutation leads to an unpaired cysteine residue and a
RT molecular weight polymorphism of a functional platelet beta 3 integrin
RT subunit. The Sra alloantigen system of GPIIb.";
RL J. Biol. Chem. 269:8439-8444(1994).
- RN [19]
RX VARIANT GTA TYR-145.
RX MEDLINE=90364410; PubMed=2392682;
RA Loftus J.C., O'Toole T.E., Plow E.F., Glass A., Frelinger A.L. III,
RA Ginsberg M.H.;
RT "A beta 3 integrin mutation abolishes ligand binding and alters
RT divalent cation-dependent conformation.";
RL Science 249:915-918(1990).
- RN [20]
RX VARIANT GLN-240.
RX MEDLINE=92156115; PubMed=1371279;
RA Bajt M.L., Ginsberg M.H., Frelinger A.L. III, Berndt M.C.,
RA Loftus J.C.;
RT "A spontaneous mutation of integrin alpha IIb beta 3 (platelet
RT glycoprotein IIb-IIa) helps define a ligand binding site.";
RL J. Biol. Chem. 267:3789-3794(1992).
- RN [21]
RX VARIANT TRP-240.
RX MEDLINE=92291320; PubMed=160206;
RA Lanza F., Stierle A., Fournier D., Morales M., Andre G., Nurdan A.T.,
RA Cazenave J.-P.;
RT "A new variant of Glanzmann's thrombasthenia (Strasbourg I).
RT Platelets with functionally defective glycoprotein IIb-IIa complexes
RT and a glycoprotein IIa 214Arg-->214Trp mutation.";
RL J. Clin. Invest. 89:1995-2004(1992).
- RN [22]
RX VARIANT STRASBOURG-1.
RX MEDLINE=93066201; PubMed=1438206;
RA Chen Y.-P., Djaffar I., Pizard D., Steiner B., Cleutaut A.-M.,
RA Caen J.P., Rosa J.-P.;
RT "Ser-752-->Pro mutation in the cytoplasmic domain of integrin beta 3
RT subunit and defective activation of platelet integrin alpha IIb beta 3
RT (glycoprotein IIb-IIa) in a variant of Glanzmann Thrombasthenia.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10169-10173(1992).
- RN [23]
RX REVIEW ON GTA VARIANTS.
RX MEDLINE=95184171; PubMed=7878622;
RA Bray P.F.;
RT "Inherited diseases of platelet glycoproteins: considerations for
RT rapid molecular characterization.";
RL Thromb. Haemost. 72:492-502(1994).
- RN [24]
RX VARIANT GTA TRP-143.
RX MEDLINE=98025892; PubMed=9376589;
RA Basani R.B., Brown D.L., Vilaire G., Bennett J.S., Poncz M.;
RT "A Leu117-->Trp mutation within the RGD-peptide cross-linking region


```

FT DISULFID 482 520 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 498 511 BY SIMILARITY.
FT DISULFID 526 531 BY SIMILARITY.
FT DISULFID 528 561 BY SIMILARITY.
FT DISULFID 533 546 BY SIMILARITY.
FT DISULFID 548 553 BY SIMILARITY.
FT DISULFID 567 572 BY SIMILARITY.
FT DISULFID 569 600 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
FT DISULFID 585 592 BY SIMILARITY.
FT DISULFID 606 611 BY SIMILARITY.
FT DISULFID 608 656 BY SIMILARITY.
FT DISULFID 613 623 BY SIMILARITY.
FT DISULFID 626 629 BY SIMILARITY.
FT DISULFID 633 642 BY SIMILARITY.
FT DISULFID 639 712 BY SIMILARITY.
FT DISULFID 660 688 BY SIMILARITY.
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 584 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOD_RES 772 772 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 784 784 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 787 AA: 86694 MW: 1570599ABC438A3 CRC64:

```

```

Query Match 92.4%: Score 3839; DB 1: Length 787;
Best Local Similarity 91.3%; Pred. No. 9, 7e-253;
Matches 694; Conservative 39; Mismatches 27; Indels 0; Gaps 0;

```

```

QY 3 NICTRGVSSCQCCCLAVPWCAMCSDALPLGSPRCDLKNLKNKCAEESIEFPVSEAR 62
DB 28 NICTRGVSSCQCCCLAVPWCAMCSDALPLGSPRCDLKNLKNKCAEESIEFPVSEAR 87
QY 63 VLEDRPLSDKSGDSSQVQVSPORALRLRPDSKNESIOVROVEDYVVDIYIMDLSY 122
DB 88 ILRLRLSLKSGSSAQIOVSPORALRLRPDSKNESIOVROVEDYVVDIYIMDLSY 147
QY 123 SMKDDMSIONLTGKTLATOMRKLTSLMRIGFGAFVDRKPSVWYISPPALEPCYDMYT 182
DB 148 SMKDDLSSTIOTLTGKTLASOMRKLTSLMRIGFGAFVDRKPSVWYISPPALEPCYDMYT 207
QY 183 TCLPMFGYKHYLTLDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVCDERIGMRDA 242
DB 208 ACIPMGYKHYLTLDQVTRFNEEVKQSVSRNRDAPEGGFDAIMQATVCDERIGMRDA 267
QY 243 SHLLVFTTAKTHIALDGRLAGIVORNDQCHVSDNHNYSASTMDYPSLGLMTEKLSQK 302
DB 268 SHLLVFTTAKTHIALDGRLAGIVORNDQCHVSDNHNYSASTMDYPSLGLMTEKLSQK 327
QY 303 NINLFAVTENVNLYONXSELIPGTVGLSKDSNNVQLIVDAVGRKRSVELEVRDL 362
DB 328 NINLFAVTENVNLYONXSELIPGTVGLSKDSNNVQLIVDAVGRKRSVELEVRDL 387
QY 363 PEBLSLSPNATCLINNEVIRGLKSCMGLKIGTVSFESIEAKVRCQDEKESFTIKPVGRK 422
DB 388 PEBLSLSPNATCLINNEVIRGLKSCMGLKIGTVSFESIEAKVRCQDEKESFTIKPVGRK 447
QY 423 DSLIYVTFDDCCACAOAQNAPNSHRNNGNTEFCGVCCGSGMGLSGQCECEEDYRPSQ 482
DB 448 DSLIYVTFDDCCACAOAQNAPNSHRNNGNTEFCGVCCGSGMGLSGQCECEEDYRPSQ 507
QY 483 QDECSREGOPVCSQREGELCGQCVCHSSDFGKITGKYECODFSVVRKSGMCGSHQC 542
DB 508 QDECSREGOPVCSQREGELCGQCVCHSSDFGKITGKYECODFSVVRKSGMCGSHQC 567
QY 543 SCGDCLSDSDMTGYCNCCTTRDTCCSSNGLLCGRKCEGSCVCIOPISYGDTCCKCP 602
DB 568 NCGDCVCDSDMTGYCNCCTTRDTCCSSNGLLCGRKCEGSCVCIOPISYGDTCCKCP 627
QY 603 TCPDACTFKKEVECKKFPKRGALHDENTCNRCRDEIESVRELKDTGDAVAVCTYKNEDD 662

```

```

DB 628 TCPDACSCKKEVECKKRNKRLHEBNTCSKRCRDIDQVELDITGNNAVCTYKNEDD 687
QY 663 CVRFRQYEDSSGKSLVVEEPECCKPDIILVLLSYGAILLIGLAALLIMKLITTH 722
DB 688 CVRFRQYEDTSGRAVLVVEEPECCKPDIILVLLSYGAILLIGLATILIMKLITTH 747
QY 723 DRKEFAKFEERAKKMDTANNPLYKATSTFTNTYNGT 762
DB 748 DRKEFAKFEERAKKMDTANNPLYKATSTFTNTYNGT 787

RESULT 3
ID ITB5_MOUSE STANDARD: PRT: 798 AA.
AC 070309; 070308; 088347.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
GN Integrin beta-5 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS BETA-5A AND BETA-5B).
RC TISSUE=Liver;
RX MEDLINE=98198405; PubMed=9531507;
RA Zhang H., Tan S.W., Lu J.;
RT "CDNA Cloning reveals two mouse betas 5 integrin transcripts distinct in
RT cytoplasmic domains as a result of alternative splicing.";
RL Biochem. J. 331:631-637(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM BETA-5A).
RC TISSUE=Brain;
RX MEDLINE=98098874; PubMed=9880508;
RA Feng X., Tettebaum S.L., Quiroz M.E., Towler D.A., Ross F.P.;
RT "Cloning of the murine betas 5 integrin subunit promoter. Identification
RT of a novel sequence mediating granulocyte-macrophage colony-
RT stimulating factor-dependent repression of betas 5 integrin gene
RT transcription.";
RL J. Biol. Chem. 274:1366-1374(1999).
RN [3]
RP FUNCTION: INTEGRIN ALPHA-V/BETA-5 IS A RECEPTOR FOR FIBRONECTIN.
CC IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.
CC -1 SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-5
CC ASSOCIATES WITH ALPHA-V.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 ALTERNATIVE PRODUCTS: 2 ISOFORMS: BETA-5A (SHOWN HERE) AND BETA-
CC 5B, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1 SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1 SIMILARITY: CONTAINS 1 WFA-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF043257; AAC40110.1; -
CC EMBL: AF043256; AAC40109.1; -
CC EMBL: AF022110; AAD08782.1; -
CC HSSP: P04355; 2MRT.
CC MGD: MGI:96614; Itgb5.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR002369; Integrin_B.
CC InterPro: IPR003659; PSI.
CC InterPro: IPR002035; WFA.
CC Pfam: PF00362; Integrin_B_1.
CC PRINTS: PR01186; INTEGRINB.
CC ProDom: PD001811; Integrin_B_1.

```

DR SMART: SM00181; EGF: 1.
 DR SMART: SM00187; INB: 1.
 DR SMART: SM00423; PSI: 1.
 DR SMART: SM00327; VWA: 1.
 DR PROSITE: PS00243; INTEGRIN_BETA: 2.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
 KM Integrin: Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 23
 FT CHAIN 1 23
 FT DOMAIN 24 798
 FT TRANSMEM 720 742
 FT DOMAIN 743 798
 FT DOMAIN 136 378
 FT DOMAIN 465 630
 FT REPEAT 465 512
 FT REPEAT 513 554
 FT REPEAT 555 593
 FT REPEAT 594 630
 FT DISULFID 28 463
 FT DISULFID 36 46
 FT DISULFID 39 75
 FT DISULFID 49 64
 FT DISULFID 202 211
 FT DISULFID 259 300
 FT DISULFID 401 413
 FT DISULFID 433 682
 FT DISULFID 461 465
 FT DISULFID 476 487
 FT DISULFID 484 522
 FT DISULFID 489 498
 FT DISULFID 500 513
 FT DISULFID 528 533
 FT DISULFID 530 563
 FT DISULFID 535 548
 FT DISULFID 550 555
 FT DISULFID 569 574
 FT DISULFID 571 602
 FT DISULFID 576 585
 FT DISULFID 587 594
 FT DISULFID 608 613
 FT DISULFID 610 657
 FT DISULFID 615 625
 FT DISULFID 628 631
 FT DISULFID 635 644
 FT DISULFID 641 714
 FT DISULFID 661 690
 FT CARBOHYD 347 347
 FT CARBOHYD 460 460
 FT CARBOHYD 479 479
 FT CARBOHYD 505 505
 FT CARBOHYD 586 586
 FT CARBOHYD 654 654
 FT CARBOHYD 705 705
 FT VARSPLIC 760 798
 FT
 FT CONFLICT 96 96
 FT CONFLICT 259 259
 FT CONFLICT 595 595
 FT SEQUENCE 798 AA: 87908 MW: 3489DB807E2688B0 CRC64:

Query Match 56.4%; Score 2343; DB 1; Length 798;
 Best Local Similarity 54.9%; Pred. No. 2.4e-151;
 Matches 428; Conservative 129; Mismatches 199; Indels 24; Gaps 11;

QY 1 GPHICTRGVSSGQCCCLAVSPKACMSDEALPLGSP-----RCDLEKNTLKNCAPESIE 55
 DB 24 GLNICTSGATSCCECLLHPKCMCKSKY-FGNPRSTSRCDLAKNIRNCGECE-IE 80
 QY 56 FVSEARVLEDRPLDKG-SGDSOVTVSPQIALRLRPDSKNSIQVROVEDPVDI 114

DB 81 SPASSTHVLRLPLSSKSSSATGSDVIOHTPOELVSLRPBEOTTFLOVQVEDYVDL 140
 QY 115 YIIMDLASYMKDDLSIONLGTIKLATQMRKLTSLNLRIGFAGVDPKPYSPMYISPEALE 174
 DB 141 YIIMDLASLMKDDLEINISLGTIKLAEMRKLTSLNLRIGFAGVDPKDISPFSY-TAPROT 199
 QY 175 NPC--YDMKTTCLPMFGKRVHVLTLTDQVTRNEEYKKOSVSRNRDAPFGCDAIMQATVC 232
 DB 200 NCIQKFLPNCVPSFGFRHLPLTRDVSFNEEYRKORVSRNRDAPFGCDAVLOAAVC 259
 QY 223 DEKIGMRDASHLVFTTDATKHTALDGRLAGIYVPOGCHVSDNYSASTMDVPSL 292
 DB 260 KETIGMRDASHLVFTTDVPHIALDGLGGLVOPHGOCHLEANEYTSNODHPSL 319
 QY 293 GLMTEKLSQKNINLFEATVENVNLQYVSELIPGTGVLSMDSNVLQILVDAVGRIR 352
 DB 320 ALIGEKLAENNNLIFATYKNHMYLKNFTALIPGTVEILHDSKNLIQILINAYSSIR 379
 QY 353 SKVELLEVDLPEELSLSFENATCLNNEVIRGLKSCMGLIGDVTVSFEAKYRGCP-QEKE 411
 DB 380 AKVELSVMDQPEDLNLFFATACODGISYPCGRKCEGLKIGDTASFEYSVEARSCPGROAA 439
 QY 412 KSEITKPYGFKNSLIVQYTFPODCACQOAAEPNSHRNNGCTPECGYRCGPGMIGSQC 471
 DB 440 QSEITLPPVGFROSLQVEAVYNCTCGSTGLEPNSARQ-SGNGTYTGIGLCECDPGLTGRC 498
 QY 472 ECSEEDYPSQODECSPPGQPVCSQORGECLCGQCVCHSSPFKCTTGCCYCCDDPSCYRY 531
 DB 499 ECGEBENSGYONLCREAEGRKPLCGRGECSCNCSCESEFGRIGYFCFCDSDSCARN 558
 QY 532 KGMCSGHQSCGQGLCDSDMTGYCNCRTTRDTQMSNGLLCSGRCKCEGSCVCTOP 591
 DB 559 KGVLCSGHECHCGCKCHAGYIGDNCNSYDVSTCAKAKDQICSDNRGRCVCGCQCTEP 618
 QY 592 GSYGPTCKPCPCPACGFEKKCEVCKKFDGALHDETCNRCYRDELES-VKELKOTGK 650
 DB 619 GAFGECECPYCPACGSKKRCVCECLLHQK-PDNOTCHHOCDEXITVVDTVKDDQ 677
 QY 651 DAVNCTYRNEDCVVRFQYEDDSGKSLIYVEEPCPKRDPILVVLVSAGVALLIGLA 710
 DB 678 EAVLFCFYTAKCCVMMFSTELPNGRSNLTVLREPEGCSAPNANMTLLAVGSIILGMA 737
 QY 711 ALLIKLLITTHDKREPAKFEERARAKMDTANPLLYEATSTT-----NTTYGCT 762
 DB 738 LLAIKLLVITTHDREPAKFOESRRARYEMASNPLYRKPISTHTVDFAFKFKNSYNGS 797
 RESULT 4
 ID ITB5_HUMAN STANDARD; PRT; 799 AA.
 AC P18084;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-5 precursor.
 GN ITGB5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymic epithelium;
 RX MEDLINE=90228356; PubMed=2328726;
 RA Ramaswamy H., Hemler M.E.;
 RT "Cloning, primary structure and properties of a novel human integrin
 beta subunit";
 RL EMBO J. 9:1561-1568(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90319111; PubMed=2371275;
 RA Suzuki S., Huang Z.S., Tanihara H.;

RT "Cloning of an integrin beta subunit exhibiting high homology with
RT integrin beta 3 subunit."
RT Proc. Natl. Acad. Sci. U.S.A. 87:5354-5358(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91009141; PubMed=2211615;
RA McLean J.W., Vestal D.J., Cheresch D.A., Bodary S.C.;
RT "cDNA sequence of the human integrin beta 5 subunit."
RL J. Biol. Chem. 265:17126-17131(1990).
CC -1- FUNCTION: INTEGRIN ALPHA-V/BETA-5 IS A RECEPTOR FOR FIBRONECTIN.
CC IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-5
CC ASSOCIATES WITH ALPHA-V.
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 VMFA-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53002; CAA37188.1; -
DR EMBL: M35011; AAA52707.1; -
DR EMBL: J05633; AAA59183.1; -
DR PIR: A35775; A35775.
DR PIR: S12534; S12534.
DR PIR: A38308; A38308.
DR PIR: S11708; S11708.
DR HSSP: P04355; 2MRT.
DR MIM: 147561; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; PSI.
DR Pfam: PF00362; Integrin_B; 1.
DR PRINTS: PRO1186; INTEGRINB.
DR PRODOM: PD001811; Integrin_B; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00187; INB; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS00243; INTEGRIN_BETA_2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 799
FT DOMAIN 24 719
FT TRANSMEM 720 742
FT DOMAIN 743 799
FT DOMAIN 136 378
FT DOMAIN 465 630
FT REPEAT 465 512
FT REPEAT 513 554
FT REPEAT 555 593
FT REPEAT 594 630
FT DISULFID 28 463
FT DISULFID 39 46
FT DISULFID 39 75
FT DISULFID 49 64
FT DISULFID 202 211
FT DISULFID 259 300
FT DISULFID 401 413
FT DISULFID 433 682
FT DISULFID 461 465
FT DISULFID 476 487
BY SIMILARITY.
INTEGRIN BETA-5.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
VMA-LIKE.
4 CYSTEINE-RICH TANDEM REPEATS.
I.
II.
III.
IV.

FT DISULFID 484 522 BY SIMILARITY.
FT DISULFID 489 498 BY SIMILARITY.
FT DISULFID 500 513 BY SIMILARITY.
FT DISULFID 528 533 BY SIMILARITY.
FT DISULFID 530 563 BY SIMILARITY.
FT DISULFID 535 548 BY SIMILARITY.
FT DISULFID 550 555 BY SIMILARITY.
FT DISULFID 569 574 BY SIMILARITY.
FT DISULFID 571 602 BY SIMILARITY.
FT DISULFID 576 585 BY SIMILARITY.
FT DISULFID 587 594 BY SIMILARITY.
FT DISULFID 608 613 BY SIMILARITY.
FT DISULFID 610 657 BY SIMILARITY.
FT DISULFID 615 625 BY SIMILARITY.
FT DISULFID 628 631 BY SIMILARITY.
FT DISULFID 635 644 BY SIMILARITY.
FT DISULFID 641 714 BY SIMILARITY.
FT DISULFID 661 690 BY SIMILARITY.
FT CARBOHYD 347 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 552 552 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 654 654 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 705 705 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 193 193 T -> A (IN REF. 2).
FT CONFLICT 645 645 L -> P (IN REF. 3).
FT CONFLICT 790 792 MISSING (IN REF. 2).
SQ SEQUENCE 799 AA; 88053 MW; D/E4727CA310512B CRC64;

Query Match 56.3%; Score 2339; DB 1; Length 799;
Best Local Similarity 55.5%; Pred. No. 4.5e-151;
Matches 433; Conservative 118; Mismatches 205; Indels 24; Gaps 11;
OY 1 GNNICTTGVSQQOCLAVSPWCAMGSDPALPLGSP-----RCDIKENLKNCAPSIE 55
DB 24 GINICTSGSATSCBCECLLIHPKCAKSKKE--DFGSPITSRCIDLRAVLKNGCGGE-IE 80
OY 56 FVSEFARVLEDRPLSDKSSGDS--QVYVSQRIARLRPPDSKNFSIQVQVQVDPYDI 114
DB 81 SPASSFHVLRSLPLSSKSGSGAGMDVQMTQDEIAVLNRPedKTFPQLQVQVDEYVDL 140
OY 115 YLMDLSYSMKDLSIONLGTATQMRKLTSLNRIQFAFVQKPSVYVISPFALE 174
DB 141 YLMDLSYSMKDLDNINSLGKTLAEEMRKLTSLNFRIGFSFVQKDISPSTY-TAPRYQT 199
OY 175 NRC--YDMKTTCLPMEGYKAVLTLLDQVTRNEVEKQSVSRNRPDAEGGFDAIMQATVC 232
DB 200 NRCIGYKLFPPNCVPSFGRHLLPLTRDVSFNEVEKQSVSRNRPDAEGGFDAVLAQAVC 259
OY 233 DEKTRWRDASHLVFTTDAKTHALDGRLAGIYQVQNDGCHVQSDNINYSATMDVPSL 292
DB 260 KEKIGWRDADHLVFTTDDVPHIALDGLGLQVPHGQCHLNEANEVYASNMQDPSL 319
OY 293 GLMTEKLSOKNINLFEAVTENVNLYONYSELIGTGYGVLSMDSNVYLIQIDVAYGKIR 352
DB 320 ALLGKTLAENNINLIFANTKKNHMYLTKNFTALIFETIYEIILDSKNIQILINAYNSIR 379
OY 353 SKVELLEVRDPELSLSFNATCLNNEVTPGLKSCMGKLTGDTVSFSIEAKVRCP-OEKE 411
DB 380 SKVELSVMDQPEDNLNLFATCCQGVSPGQKCEGLKIGTASFEVSLERASCPSSRTE 439
OY 412 KSFTIKPVGFDLSLIYVTPDCDACAQAAPRNSHRNNGNGTECEGVCRCRGPWLSQC 471
DB 440 HVFALRPVGFPRSLSEVGTYNCTGCGSYGLBPNARC-NSGTYVCCICCESPYLTGRC 498
OY 472 ECSEEDYPPSODECSPEPGQPVCSOREGCLCGQCVCHSPFGKITKCYKCEDFQSVRY 531
DB 499 ECQDGENSVYQNLCREAGRPGLSGKDCSCNCSCEFSFEGKITYPGFCEDNFSQARN 558
OY 532 KGMCSGHGQSCGDCCLCDSDMTGYCNCCTRTDTCMSSNGLCSGRGKCEGSCVCIOP 591

```

Db 559 GVCISGHCCECHCECHAGYICDNCSTIDSTCRGRDQICSEHGCLCQCCTEP 618
Qy 592 GSYDTECEKPTCPDCTPFKECEVCKRFGALHDETCNRCDIES-VKELDTGK 650
Db 619 GAFGECEKPTCPDCTPFKECEVCKRFGALHDETCNRCDIES-VKELDTGK 677
Qy 651 DAVNCTKNEDECVRRQYVEDSGKSLVVEEPCPKGPDILLVSLVGAALLIGLA 710
Db 678 EAVLCFYTAKDCVMFTYVELPSGKSNLVLRPEECGNPNMTILLAVGSLILVGLA 737
Qy 711 ALLIMKLLITIDHKEFEKFEERAKMPTANPLKATSTFT-----NITKRT 762
Db 738 LLAIMKLLITIDHKEFEKFEERAKMPTANPLKATSTFT-----NITKRT 797

RESULT 5
ID ITB6.MOUSE STANDARD; PRT; 787 AA.
AC 092079;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-6 precursor.
GN ITGB6.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=20547423; PubMed=11095652;
RA Arend L.J., Smart A.M., Briggs J.P.;
RT Mouse beta(6) integrin sequence, pattern of expression, and role in
RT kidney development."
RL J. Am. Soc. Nephrol. 11:2297-2305(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-V/BETA-6 IS A RECEPTOR FOR FIBRONECTIN
CC AND CYTOTACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS
CC (BY SIMILARITY).
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-6
CC ASSOCIATES WITH ALPHA-V (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF115376; AAD17212.1; -
CC MGD: MGI:96615; Itgb6.
CC DR InterPro: IPR000561; EGF-like.
CC DR InterPro: IPR002369; Integrin_B.
CC DR InterPro: IPR001169; Integrin_beta_C.
CC DR InterPro: IPR003659; PSI.
CC DR pfam: PF00362; Integrin_B.1.
CC DR PRINTS: PR01186; INTEGRINB.
CC DR ProDom: PD001811; Integrin_B.1.
CC DR SMART: SM00001; EGF_like.1.
CC DR SMART: SM00187; INB.1.
CC DR SMART: SM00423; PSI.1.
CC DR SMART: SM00327; VWFA.1.
CC DR PROSITE: PS00243; INTEGRIN_BETA.2.
CC DR PROSITE: PS00032; EGF_1; UNKNOWN_2.
CC DR PROSITE: PS01186; EGF_2; UNKNOWN_1.1.
CC DR Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
CC Repeat; Signal.

```

```

FT SIGNAL 1 21
FT CHAIN 22 787
FT DOMAIN 22 706
FT TRANSMEM 707 729
FT DOMAIN 730 787
FT DOMAIN 131 371
FT DOMAIN 456 619
FT REPEAT 456 501
FT REPEAT 502 543
FT REPEAT 544 582
FT REPEAT 583 619
FT DISULFD 23 454
FT DISULFD 31 41
FT DISULFD 34 70
FT DISULFD 44 59
FT DISULFD 197 204
FT DISULFD 252 293
FT DISULFD 394 406
FT DISULFD 426 669
FT DISULFD 452 456
FT DISULFD 467 479
FT DISULFD 476 511
FT DISULFD 481 490
FT DISULFD 492 502
FT DISULFD 517 522
FT DISULFD 519 552
FT DISULFD 524 537
FT DISULFD 539 544
FT DISULFD 558 563
FT DISULFD 560 591
FT DISULFD 565 574
FT DISULFD 576 583
FT DISULFD 597 602
FT DISULFD 599 645
FT DISULFD 604 614
FT DISULFD 617 620
FT DISULFD 624 633
FT DISULFD 630 701
FT DISULFD 649 677
FT CARBOHYD 48 48
FT CARBOHYD 97 97
FT CARBOHYD 260 260
FT CARBOHYD 387 387
FT CARBOHYD 418 418
FT CARBOHYD 463 463
FT CARBOHYD 471 471
FT CARBOHYD 541 541
FT CARBOHYD 575 575
SQ SEQUENCE 787 AA; 86041 MW; C6438C6F1E6B7FBD CMC64;

Query Match 52.4%; Score 2177; DB 1; Length 787;
Best Local Similarity 51.6%; Pred. No. 4,2e-140;
Matches 392; Conservative 124; Mismatches 232; InDels 12; Gaps 7;

Qy 5 CTTREYSSCQCLAVSPMCACMSDE---ALPLGSRCLKENLKDNCAPSIIEPVSSEA 61
Db 23 CAGGAECSCDCLLPGPCACMSQENFTHLSGAGEKCDTPANLLAKGQLPFIENPVSRI 82
Qy 62 RYLEDPLSDKSGDSQVOTGVSPORIALRLRPDSDKSFNSIQVQVQVEDYPPVYLDMS 121
Db 83 EYLQKKPLSYGQKNSDIOVLAPOSILVKLRPGREGQLOVQVQKOTEDYPPVLYLDMS 142
Qy 122 YSMKDLMSLONLGTKLATQMKLTSNLRIGFAGVDPVSPYVYISPEALNEPCYDMK 181
Db 143 ASMDDLNLTIKELGSRLLAKESKLTNSRFLGFSVEKVPSPFK-TTPEITNPPCSSIP 201
Qy 182 TTCLEMPGYKHYVLTLDQVTRFNEVKKQSVSRNDAPGEGFALIMQATVDEKIGRND 241
Db 202 YFCLETFGFKHLLPLTDAERFNEIVRKQISANIDPPEGGFDAIMQAAVCKEIKGRND 261
Qy 242 ASHLLVFTTDAKTHALDGLRAGIVQPNDCGHGSDNHYASATJMDVPSLGLMTEKLSQ 301

```


FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 788 AA: 85935 MW: EDBYD533EC4C8C4D CRC64;

Query Match 52.3%; Score 2170.5; DB 1; Length 788;
 Best Local Similarity 50.9%; Pred. No. 1.2e-139;
 Matches 387; Conservative 134; Mismatches 228; Indels 11; Gaps 7;

5 CTRGVSSCCOCCAVSPMCASDEAL-PLG-SPKCDLKENLKNCAPESIEFVSEA 61
 23 CALGAATCTCDLLIGQCACMAQENTPHSVGGERCDTPANLLAGCQINFLENVSQV 82
 62 RVEDRLSKSGSDSQVTPQRTALRLRPDSKNFSIOVQVEDYVVDIYIMDS 121
 83 ELKKNPLSYGRKNSNDIQAIPQSLILKLRGGAQTQVHVROTEDYVVDIYIMDS 142
 122 YSMKDDLMSTONLGTAKATQMRKLTSLRIGFAGYDKPSPMYTISPAPALENPEYDK 181
 143 ASMDDLNLTKEGLSKSKMSKLTSLRIGFAGYDKPSPMYTISPAPALENPEYDK 201
 182 TTCLPMFGYKVVLTLDQVTFRENEEYKQSVSHNRDAPESGPDALMQATVCEDEKIMRD 241
 202 YFLCFPGFKHILPLTNDARFNEIYKNOISANIDTPEGGFALMQAVCKRCKIOMRD 261
 242 ASHLVPTTAKTHIALDGLAIVQPNQCCVHSGSDNHTSASTMDYSLGLMTEKLSQ 301
 262 SLHLVFSVSDSHFEGDSKLAGIIVPNDGLCHLDSKNEYSMTVLEYPTIGLIDKLVQ 321
 302 KNINLFAVENVNYLYQNTSELIPGTGYVLSMSSNVLQTLVDYAGKIRKSEVEVLD 361
 322 NNVLILFAVYQEVNHLKENTAKILPGATVGLLOKDSNLTQILITSYEEELRSVELEVLG 381
 362 LPELISLENATGLNNEVPLGLKSCMGLIGTIVSEISAKVYKGPQEKESFTIPVGF 421
 382 DTGLNLSTFALINNGTLEFQHKCKSHMKVGDYASFSVYVNIPIHC-ERRSRHIIKPVGL 440
 422 KDLIVQVFPDCCACAOAEPNSHRNNGNGTFEGGVCRCGGLGSGCESEEDYRPS 481
 441 GDALVLEVSPECNCDQKEVEVNSKCHHGNSFCQGVACAFHGMGPCEGED---ML 497
 482 QODECSPREGQPVCSORGECLGQCYCHSSDFKIKTGKCECDPFSCVYKKGEMSGHQ 541
 498 STDSCKEADHPSCSGRGDCYCGQCICHSPIGNITGPYCQCDNFSVHRKGLLCGGND 557
 542 CSGGCDLSDWMTGYCNCCTRTDTCMSNGLLSCRGKCEGSCYCIOPGSGYDCEK 601
 558 CDSGECVCSNGMGTCTTSTDCVSEIDVGLSCRGDCYCGKCTNPGASGPTCEK 617
 602 PTCPRDCTKKEVECKKDRGLADHENTCNRYCRDEISVLEKDTGKD-AVNCTYKNE 660
 618 PTCGDCNSKRSCTECHLSAQAQARE--CVDCKLAGATISEEDFESKDSVCSLQGE 675
 661 DDCVVFQYEDSGKSLIYVEEPCPGPDILVLSVMATILIGLALLIKLLIT 720
 676 NEGLITFLTTDNEGTTIHSINEKDCPPNIPMILGVALIIGVLLCTIKLLVS 735
 721 HDRKEFAFEERARAKWDANPLYEATSTFTNITYR 760
 736 FHDRKEVAFKFAERSKAKWOTGNPLYRGSTSTFKNVYK 775

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Integrin beta-5 (Fragment).
 GN ITGB5.
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 OC NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94040831; PubMed=8224922;
 RA Shoji M., Hayzer D.J., Kim T.M., Runge M.S., Hanson S.R.;
 RT "Human and baboon Integrin beta 5 subunit-encoding mRNAs have
 RT alternative polyadenylation sites.";
 RL Gene 133:307-308(1993)
 CC -1 FUNCTION: INTEGRIN ALPHA-V/BETA-5 IS A RECEPTOR FOR FIBRONECTIN.
 CC -1 RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.
 CC -1 SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-5
 CC -1 ASSOCIATES WITH ALPHA-V.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS.
 CC -1 SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1 SIMILARITY: CONTAINS 1 WFMA-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L12231; AAA16866.1; -
 DR HSEF: P04355; 2MRP; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR002035; WFMA.
 DR Pfam: PF00362; Integrin_B; 1.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00327; VMA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA_2.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Extracellular matrix; Cytoskeleton.
 FT NON_TER 1 1
 FT DOMAIN <1 575
 FT TRANSMEM 576 598
 FT DOMAIN 599 655
 FT DOMAIN <1 234
 FT DOMAIN 457 621
 FT REPEAT 321 368
 FT REPEAT 369 410
 FT REPEAT 411 449
 FT REPEAT 450 486
 FT CARBOHYD 203 203
 FT CARBOHYD 316 316
 FT CARBOHYD 408 408
 FT CARBOHYD 442 442
 FT CARBOHYD 510 510
 FT CARBOHYD 561 561
 SQ SEQUENCE 655 AA: 72466 MW: BAC3JA159CBE1596 CRC64;

Query Match 47.9%; Score 1989.5; DB 1; Length 655;
 Best Local Similarity 55.5%; Pred. No. 1.7e-127;
 Matches 364; Conservative 102; Mismatches 175; Indels 15; Gaps 7;

QY 119 DISVSMKDDLSIONLGTAKLTQNRKLTSLNRIGFAGVDRKPSVPMYISPPLEALNCP- 177
 1 DLSLSMKDDLDITRLMGLKLAEMKRLTSLNFRLLGSGFVDDKDISPFSY-TAPRYOTNPCI 59
 QY 178 -YDMKTTCLPMGYHNVLTLDQVTRFNEEYKOSVSNRNPAPREGGPAIMQAVCDCKI 236
 60 GYKLPNCVPSTGFRRHLPLDTRDVSFEYKORQVSNRNPAPREGGPAIMQAVCDCKI 119
 QY 237 GWRNDASHLIVETTDKATHIALDGLRAGIYQPNQDQCHVSGDNHYSASTMDPSLGMT 296
 120 GMRKALHLIVETTDVPHIALDGLGLQPHDQCHLNEANEYTAQNQMDYPSLALG 179
 QY 297 EKLQSKNINLIPAVENYVNLQYNSLIPCTGVVLSMSSNVQLIYDAVAKIRKVE 356
 180 EKLANNINLIPAVYKNMYLMYKNTALIPCTVEIILDGSKNIIQLITINYNISRSKVE 239
 QY 357 LEVRDLPEELISEFNATCLNNEVIRPGLKSGMGLTGDVPSIIEAKVGC-0EKEKST 415
 240 LSVMDQEDNLFTATQDQGVSTPGQKCEGLKIGDTASEVSEVSEANSCSTRHEHYFA 299
 QY 416 IKPVGKSLIVQVTFDCCACAOAEPSNRHRCNNGNGTFCQVGCRCGPMGLGSCQCESE 475
 300 LQPVCRDLSLEVGYVNTCCGCVGLFEPNSARC-SGTGYVGCLECECPGLYGRKCEQD 358
 QY 476 EYRSQDQECSPREGQVYCSQRGLCGQCVCHSDRGKITGKCECDSDSVYRKQEM 535
 359 GNNHVSYNLCLDTEGKPLCSGRCSCNQCSCFESFEKTIYGPCECDNFSCANKGV 418
 QY 536 CSGHGCSCGDCDLDSDMTGYCNCNCTRTDQMSNGLCSGRGKCEGSCVCIQPGSYG 595
 419 CGSHGCHGCECKHAGYIGDNCNSTDITCRGRDGLICSERGHCLCGQCGCTPRPAG 478
 QY 596 DTCEKCPCTPDACITRKECEVCKKFRDGLADHENTCNRYCDEITS-VKELDKGKAVN 654
 479 EKCEKCPCTPDACITRKECEVCKKFRDGLADHENTCNRYCDEITS-VKELDKGKAVN 537
 QY 655 CTYKNEKDCVRFQYEDDSGKSLIYVEEPCRPDILVLLSVMAIILLIGLALILI 714
 538 CTYKNAKDCVRFQYEDDSGKSLIYVEEPCRPDILVLLSVMAIILLIGLALILI 597
 QY 715 WLLITTHDRKEFAFEERARAKWDANPLYEKATSTFT-----NITYRGT 762
 598 WLLITTHDRKEFAFEERARAKWDANPLYEKATSTFT-----NITYRGT 653
 Db
 RESULT 8
 ITBL_CHICK STANDARD: PRT: 803 AA.
 AC P07228;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-1 precursor (CSAR antigen) (UG22 antigen) (RGD-
 receptor).
 GN ITGB1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic fibroblast;
 RX MEDLINE=86245073; PubMed=3487386;
 RA Tamkun J.W., Desimone D.W., Fonda D., Patel R.S., Buck C.,
 Horwitz A.F., Hynes R.O.;
 RT "Structure of Integrin, a glycoprotein involved in the transmembrane
 RT linkage between fibronectin and actin.";
 RL Cell 46:271-282(1986).
 CC -I- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-
 CC 1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
 CC 1/BETA-1 AND BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED

CC CC SEQUENCE G-P-D-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
 CC 3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-6/BETA-1, ALPHA-
 CC 10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
 CC FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
 CC THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
 CC INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN
 CC ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1
 CC ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
 CC FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-T-D-S IN VCAM1. INTEGRIN
 CC ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTACTIN AND
 CC OSTEOPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN
 CC CYTOTACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
 CC AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
 CC VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
 CC WIDE ARRAY OF LIGANDS.
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
 CC ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-6, ALPHA-
 CC 5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
 CC ALPHA-V.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -I- SIMILARITY: CONTAINS 1 VFMA-LIKE DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M14049; AAA48926.1; -
 CC PIR: A23947; ITCB3.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; VFMA.
 DR Pfam: PF00362; Integrin_B_1.
 DR PRINTS: PR01186; INTEGRINB.
 DR PRODOM: PD001811; Integrin_B_1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VMA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Phosphorylation; Signal.
 FT SIGNAL 1 24
 FT CHAIN 1 803
 FT DOMAIN 25 733 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 734 756 POTENTIAL.
 FT DOMAIN 757 803 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 144 382 VFMA-LIKE.
 FT DOMAIN 471 640 4 CYSTEINE-RICH TANDEM REPEATS.
 FT REPEAT 471 520 I.
 FT REPEAT 521 564 II.
 FT REPEAT 565 603 III.
 FT REPEAT 604 640 IV.
 FT DISULFID 31 469 BY SIMILARITY.
 FT DISULFID 39 49 BY SIMILARITY.
 FT DISULFID 42 79 BY SIMILARITY.
 FT DISULFID 52 68 BY SIMILARITY.
 FT DISULFID 211 217 BY SIMILARITY.
 FT DISULFID 265 305 BY SIMILARITY.
 FT DISULFID 405 419 BY SIMILARITY.
 FT DISULFID 439 469 BY SIMILARITY.
 FT DISULFID 467 471 BY SIMILARITY.
 FT DISULFID 482 494 BY SIMILARITY.
 FT DISULFID 491 530 BY SIMILARITY.
 FT DISULFID 496 505 BY SIMILARITY.
 FT DISULFID 507 521 BY SIMILARITY.
 FT DISULFID 536 541 BY SIMILARITY.


```

FT DISULFID 538 573 BY SIMILARITY.
FT DISULFID 543 558 BY SIMILARITY.
FT DISULFID 560 565 BY SIMILARITY.
FT DISULFID 579 584 BY SIMILARITY.
FT DISULFID 581 612 BY SIMILARITY.
FT DISULFID 586 595 BY SIMILARITY.
FT DISULFID 597 604 BY SIMILARITY.
FT DISULFID 618 623 BY SIMILARITY.
FT DISULFID 620 666 BY SIMILARITY.
FT DISULFID 625 635 BY SIMILARITY.
FT DISULFID 638 641 BY SIMILARITY.
FT DISULFID 645 654 BY SIMILARITY.
FT DISULFID 651 728 BY SIMILARITY.
FT DISULFID 670 704 BY SIMILARITY.
FT MOD_RES 25 25 BLOCKED.
FT MOD_RES 788 788 PHOSPHORYLATION (BY TYR-KINASES)
(POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 589 589 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 803 AA; 88553 MW; 2F6FEFCDF2C80457 CRC64;

```

```

Query Match 44.5%; Score 1846.5; DB 1; Length 803;
Best Local Similarity 45.1%; Pred. No. 1, 1e-117;
Matches 353; Conservative 136; Mismatches 266; Indels 27; Gaps 15;

```

```

QY 1 GPNICTRGVSSCCOCLAVSPMCAMC--SDEALPLGSP---RCOLKENLMDNCAPESIEF 56
D 27 GGSPICIANAKKSCGCEICIOAGPCNCMKCTDFLOGEPTSRCDLALAKSKCEPDIDEN 86
QY 57 PVSAARVLEDRPLSDKSGSSQ-----VTVQSFORTIALRLRPDSKNFSIOVROVEDYP 111
D 87 PRGSRVLEDRREYVNRKIGAEKRLKPEALITQIOPOKLVLRGVEPOTFSLKFRRAEDYP 146
QY 112 VDIYVLMDSYMSKDDLMSTONICTKATATMKRLTSLRGFGAFYDKPSPIWTSPE 171
D 147 IDLYVLMDSYMSKDDLMSTONICTKATATMKRLTSLRGFGAFYDKPSPIWTSPE 171
QY 172 ALENPICYDMKTCLPMFGYKHLVLTLDQVTRFNEEVYKQSVSRNRDPAPEGCFPAIMQATV 231
D 207 -LRNPGCG-DQNCISPESTYKKNVLSLSEGNKFNELVNGKOHISGNLDSPEGCFPAIMQAV 264
QY 232 CDEKIGRNDASHILVFTTDAKTHIALDGLAGIYVQNDGCHVSGDNHYSASTNDYPS 291
D 265 CGGDIGIRN-VTRILVFSTAGFHFAGDGKLGIVLPNDKCHL-ENNMVYTMSHYDYS 322
QY 292 LGLMTEKLSOKNINLIFAVENVVNLVYONSELIPGTIVYVSMDSNNVQLIVDAKGI 351
D 323 IAHVQKLSNNIOTIFATVEFOAVYKELKNLIPKSAVGLTSSNSNVYQLIIDVAYNSL 382
QY 352 RSKVLEVRDLPEELSLFNATCLN--NEVYPLKSCMGILKIDTYSFSTEAKVRKCPQ 409
D 383 SSEVYIENSKLPKEVITISKYCKNGVNDIQEDGRKNSISIDEVRFELNVTANCPK 442
QY 410 -KKSFTIKVYGRKDSLIYVTFDCCACQAOAEPRSHKNGNNGTFEGVGCRCGGMUG 468
D 443 GQNETTIKIKLFTVEEHLHLOFCCLQSEEPSPACHPDNGFEFCGACGRCNGRIG 502
QY 469 SOCESEEDYRPSQDE--CSPRGQPVCSORGECLCGCCHSDGK--ITGKYCEGD 525
D 503 RLCESGIDEVNSDMAYKCRENSTEICSNGECIGQCVCKKRENTNEVYSKICECN 562
QY 526 FSCVRYKGEKSGHGQSCGDCCLDSDMWGYCNCCTTRTDGCMSSNGLLCGRKCEGS 585

```

```

DB 563 FNCDRSNGLLCGGNGLCKCVCYCFPNFTGSACDCLDTTPCAWAGNGLCNGNGTCEGT 622
QY 586 CVCIOGSGYDDECEKPTCPDCTAFKKEVECKKFPGRGALHDENTNRC-----RDEIS 641
D 623 CNETDPRFQGPTECKMOTCTGVCARHKDCVCAFAFKG--EKKETCSQDCMHFNMTRVES 680
QY 642 VKEL-KDTGKDAV-NCTYKNEDDCVVRFOYEDDSGKSLIYVEEPCPKPDILVLLS 699
D 681 RGLLPQVPHDPLSHCKENDVGCWFFYTVSNVSNASVHVYETPCSGPDILYVNG 740
QY 700 VMCALLIGLALILKLLTTHDRKEPAFEERARANDVANNDELVEATSTFTNTTY 759
D 741 VVAGIVLIGLALLILKLLIHDREFAKFEKEMKNAWDGENDPIYKSAVTYVNPAY 800
QY 760 RG 761
D 801 EG 802
DB 801 EG 802
RESULT 9
ID ITBL_FELCA STANDARD; PRT; 798 AA.
AC P53713:
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Integrin beta-1 precursor (fibronectin receptor beta subunit)
DE (CD29) (Integrin VLA-4 beta subunit).
GN ITGB1.
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fells.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Willett B.J.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-
CC 1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
CC 1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED
CC SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
CC 3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-6/BETA-1, ALPHA-
CC 10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
CC FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
CC THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
CC INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINGEN. INTEGRIN
CC ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1
CC ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
CC FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-T-D-S IN VCAM1. INTEGRIN
CC ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOACTIN AND
CC OSTEOPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN
CC CYTOACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
CC AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
CC VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
CC WIDE ARRAY OF LIGANDS.
CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
CC ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
CC 5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
CC ALPHA-V.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC - SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC DR EMBL; U27351; AAC19407.1; -

```

DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta.C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR02035; VWFA.
 DR Pfam: PF00362; Integrin_B.1.
 DR PRINTS: PR01186; INTEGRINB.
 DR PRODOM: PD001811; Integrin_B.1.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWFA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA.3.
 DR PROSITE: PS00022; EGF_1; UNKNOWN.2.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal; Phosphorylation.
 FT CHAIN 1 20 BY SIMILARITY.
 FT DOMAIN 21 798 INTEGRIN BETA-1.
 FT TRANSMEM 21 728 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 729 751 POTENTIAL.
 FT DOMAIN 752 798 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 140 378 VWFA-LIKE.
 FT REPEAT 466 635 4 CYSTEINE-RICH TANDEM REPEATS.
 FT REPEAT 466 515 1.
 FT REPEAT 516 559 1.
 FT REPEAT 560 598 1.
 FT REPEAT 599 635 1.
 FT DISULFID 27 464 IV.
 FT DISULFID 35 45 BY SIMILARITY.
 FT DISULFID 38 75 BY SIMILARITY.
 FT DISULFID 48 64 BY SIMILARITY.
 FT DISULFID 207 213 BY SIMILARITY.
 FT DISULFID 261 301 BY SIMILARITY.
 FT DISULFID 401 415 BY SIMILARITY.
 FT DISULFID 435 691 BY SIMILARITY.
 FT DISULFID 462 466 BY SIMILARITY.
 FT DISULFID 477 489 BY SIMILARITY.
 FT DISULFID 486 525 BY SIMILARITY.
 FT DISULFID 491 500 BY SIMILARITY.
 FT DISULFID 502 516 BY SIMILARITY.
 FT DISULFID 531 536 BY SIMILARITY.
 FT DISULFID 533 568 BY SIMILARITY.
 FT DISULFID 538 553 BY SIMILARITY.
 FT DISULFID 555 560 BY SIMILARITY.
 FT DISULFID 574 579 BY SIMILARITY.
 FT DISULFID 576 607 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 FT DISULFID 592 599 BY SIMILARITY.
 FT DISULFID 613 618 BY SIMILARITY.
 FT DISULFID 615 661 BY SIMILARITY.
 FT DISULFID 620 630 BY SIMILARITY.
 FT DISULFID 633 636 BY SIMILARITY.
 FT DISULFID 640 649 BY SIMILARITY.
 FT DISULFID 646 723 BY SIMILARITY.
 FT DISULFID 665 699 BY SIMILARITY.
 FT MOD_RES 783 783 PHOSPHORYLATION (BY SIMILARITY).
 FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 584 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 798 AA; 88092 MW; 2A1C3871046D838 CRC64;

Query Match 44.1%; Score 1833; DB 1; Length 798;
 Best Local Similarity 44.6%; Pred. No. 8.7e-117;

Matches 348; Conservative 136; Mismatches 268; Indels 28; Gaps 13;
 QY 3 NICTTGAVSSCQCLAVSPWCACSD-----ELPLPGSPRCDLKELLKDNCPASIEPP 57
 DB 25 NRIKLANAKSGGICGAGPNCWCNVSTLQESMP-SARCDLLEALKKGGCHPDIEPP 83
 QY 58 VSEARVLEDPRLSDKSGDSSQ-----VTVQSPRIALRLPDDSKNFSIOVRQVEDPY 112
 DB 84 RGSKDYKKNKNVNNRSGGTKEKLPEDINQIQOQVLQLRSEGPFTLKFRAADYPI 143
 QY 113 DIYIMDLSTSMKDLMSIONLCTKATQMRKLTSLRGFAVYKPPSPMYISPPA 172
 DB 144 DLYIMDLSTSMKDLLENVSLCTDLINERRTTSDFRIGFSGEYKVPYISTTPAK- 202
 QY 173 LENDPCYMTTCLPMGYKHVLTLDQVTFENVEVKOSYSRRNDAPEGGFDAIMQATVC 232
 DB 203 LRNFC-TSEONCTSPSYKAVLSLTDKGFENLVGKRISGULDSPEGFDAIMQAVNC 261
 QY 233 DEKIGWRNDASHLVFTTDAKTHIALDGLAGIVQPNDCGHVSDNNHYSASTMDYPSL 292
 DB 262 GSLIGWRN-VTRLVFSTDAGFHAGDGKLGIVLPNDGCHLEND-VYTMSHYDYPSI 319
 QY 293 GLMTERLSQKNINLIPAVTENYVNIQNSBELIPGTVGVLSMDSNVLIQIYDAYGXR 352
 DB 320 AHIWQKLSENNIOPIFAVTEEPQYKELNLIIPKSAVGLSANSSNVLIQIYDAYNSLS 379
 QY 353 SKVELEVRLPEELSLSFNATCLN-NEVYIPGLSKMGKIGDVSFSEAKVRCGPQEK 410
 DB 380 SEVLENSKIPLEGVITISYKCKNGVNGTEENRCKNSISIGDEVFELISTANKCPNNK 439
 QY 411 EKSEFTIPVGRKOSLIQVTFDDCACAQAEPSNRHCNNNGKTFEGCYVRCRGPMLGSG 470
 DB 440 SEIKIKIPLFTEVEVETIILQIFCECEQNEGIPSSPCHEGNSPFCGACRKEGVGRH 499
 QY 471 CEGSEEDYRPSQODE-OSPREGQPVCSORGECLGQCVCHSPDFGK--ITYGKCECDPS 527
 DB 500 CECSTDEVNSMDATYRKENSSEIGCSNNECEVCGQCVCKRNTNMTISGKCECDNEN 559
 QY 528 CVRYKMGSGHGGSCGDCGLCDSDWTGYCNCCTRTTCMSSNGILCSGRGCEGSGCY 587
 DB 560 CDRSNGILCGNGVCRCVCECNPNYGSACDCLDITTSICMATNGOICNRGICEGCG 619
 QY 588 CIOPGSYGDCCEKCPDCTCFKCEVECKKDRGALHDEMCNRCYR---DELESYK 643
 DB 620 CTDRKFGPCEMCQTLGVAEHKEVCQCRAFKGGKD--TCAQFCSHFNITKENRD 677
 QY 644 ELKDTGR-DAVNCTYNEEDCVRFQYDSDSGSLIYVEREPCPKGPDILVLLSVN 701
 DB 678 KLPPGQVDPLSHCKEKDQVDCWFYFTYSVNGNEALVHVVEPCTGPDIPIVAGVY 737
 QY 702 GAILLIGLALLIKLLITTHDKREPAKFEERARAKMDTANNPLKFAISTNTITYRG 761
 DB 738 AGIVILIGLALLIKLLIMTIHDTREFAKFEKKNAMKMDGENPIYKSAVTVVNPKEYG 797
 RESULT 10
 ID ITBI_HUMAN STANDARD; PRT; 798 AA.
 AC P05556; P78466; P78467; Q13089; Q14647; Q13090; Q13212; Q13091;
 AC Q14622;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
 DE (CD29) (Integrin VLA-4 beta subunit).
 GN ITGB1 OR FNBR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN 111
 RP SEQUENCE FROM N.A. (ISOFORM BETA-1A).
 RC TISSUE=Placenta;

RX MEDLINE=8807843; PubMed=2958481;
 RA Argaves W.S., Suzuki S., Arai H., Thompson K., Pletschbacher M.D.,
 RA Rusoltsch E.;
 RT "Amino acid sequence of the human fibronectin receptor.";
 RL J. Cell Biol. 105:1183-1190(1987).
 RN [2]
 RX SEQUENCE OF 717-757 FROM N.A. AND CHARACTERIZATION OF BETA-1B.
 RX MEDLINE=93209984; PubMed=7681433.
 RA Balzac F., Belkin A.M., Koteliarsky V.E., Balabanov Y.V., Altruda F.,
 RA Silengo L., Tarone G.;
 RT "Expression and functional analysis of a cytoplasmic domain variant of
 the beta 1 integrin subunit.";
 RL J. Cell Biol. 121:171-176(1993).
 RN [3]
 RP SEQUENCE OF 717-757 FROM N.A. AND FUNCTION.
 RX MEDLINE=95014744; PubMed=7523423;
 RA Balzac F., Retta S.F., Albini A., Melchiorri A., Koteliarsky V.E.,
 RA Geuna M., Silengo L., Tarone G.;
 RT "Expression of beta 1b integrin isoform in CHO cells results in a
 dominant negative effect on cell adhesion and motility.";
 RL J. Cell Biol. 127:557-565(1994).
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS BETA-1A; BETA-1B; BETA-1C AND
 BETA-1D).
 RX TISSUE-SKELETAL MUSCLE;
 RX MEDLINE=95398646; PubMed=7545396;
 RA Zhidkova N.I., Belkin A.M., Mayne R.;
 RT "Novel isoform of beta 1 integrin expressed in skeletal and cardiac
 muscle.";
 RL Biochem. Biophys. Res. Commun. 214:279-285(1995).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1B).
 RX TISSUE-SKELETAL MUSCLE;
 RX MEDLINE=91065359; PubMed=2249781;
 RA Altruda F., Cervella P., Tarone G., Balzac F., Stefanuto G.,
 RA Silengo L.;
 RT "A human integrin beta 1 subunit with a unique cytoplasmic domain
 generated by alternative mRNA processing.";
 RL Gene 95:261-266(1990).
 RN [6]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1D).
 RX MEDLINE=95377431; PubMed=7544298;
 RA van der Flier A., Kulkman I., Baudoin C., van der Neut R.,
 RA Sonnenberg A.;
 RT "A novel beta 1 integrin isoform produced by alternative splicing:
 unique expression in cardiac and skeletal muscle.";
 RL FEBS Lett. 369:340-344(1995).
 RN [7]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS BETA-1A AND BETA-1C).
 RX TISSUE-CERVICAL CARCINOMA;
 RX MEDLINE=92202279; PubMed=1551917;
 RA Langino L.R., Rusoltsch E.;
 RT "An alternative form of the integrin beta 1 subunit with a variant
 cytoplasmic domain.";
 RL J. Biol. Chem. 267:7116-7120(1992).
 RN [8]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1C-2).
 RX MEDLINE=98161805; PubMed=9494004;
 RA Slieng G., Fressler R., Johansson S.;
 RT "Identification of beta1c-2, a novel variant of the integrin beta1
 subunit generated by utilization of an alternative splice acceptor
 site in exon C.";
 RL Biochem. J. 330:1255-1263(1998).
 CC -1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1
 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
 1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED
 SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
 3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-
 10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
 FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
 THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
 INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN
 ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1

CC ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
 CC FOR VCAM1. IT RECOGNIZES THE SEQUENCE O-I-D-S IN VCAM1. INTEGRIN
 CC ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTACTIN AND
 CC OSTROPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN
 CC CYTOTACTIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
 CC AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
 CC VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
 CC WIDE ARRAY OF LIGANDS. ISOFORM BETA-1B INTERFERES WITH ISOFORM
 CC BETA-1A RESULTING IN A DOMINANT NEGATIVE EFFECT ON CELL ADHESION
 CC AND MIGRATION (IN VITRO).
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
 CC ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
 CC 5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
 CC ALPHA-V.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ISOFORM BETA-1B
 CC DOES NOT LOCALIZE TO FOCAL ADHESIONS.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: BETA-1A (SHOWN HERE), BETA-1B,
 CC BETA-1C, BETA-1C-2 AND BETA-1D; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: ISOFORM BETA-1A IS WIDELY EXPRESSED, OTHER
 CC ISOFORMS ARE GENERALLY COEXPRESSED WITH A MORE RESTRICTED
 CC DISTRIBUTION. ISOFORM BETA-1B IS EXPRESSED IN SKIN, LIVER,
 CC SKELETAL MUSCLE, CARDIAC MUSCLE, PLACENTA, UMBILICAL VEIN
 CC ENDOTHELIAL CELLS, NEUROBLASTOMA CELLS, LYMPHOMA CELLS, HEPATOMA
 CC CELLS AND ASTROCYTOMA CELLS. ISOFORMS BETA-1C AND BETA-1C-2 ARE
 CC EXPRESSED IN MUSCLE, KIDNEY, LIVER, PLACENTA, CERVICAL EPITHELIUM,
 CC UMBILICAL VEIN ENDOTHELIAL CELLS, FIBROBLAST CELLS, EMBRYONAL
 CC KIDNEY CELLS, PLATELETS AND SEVERAL BLOOD CELL LINES. ISOFORM
 CC BETA-C-2, RATHER THAN ISOFORM BETA-1C, IS SELECTIVELY EXPRESSED IN
 CC PRIMARY T-CELLS. ISOFORM BETA-1C IS EXPRESSED IN NONPROLIFERATING
 CC AND DIFFERENTIATED PROSTATE GLAND EPITHELIAL CELLS. ISOFORM BETA-
 CC 1D IS EXPRESSED SPECIFICALLY IN STRIATED MUSCLE (SKELETAL AND
 CC CARDIAC MUSCLE).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTRAINS 1 WFA-LIKE DOMAIN.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD29 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd29.htm".
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X07979; CAA30790.1; -
 CC EMBL: U33882; AAA79835.1; -
 CC EMBL: U33879; AAA79835.1; JOINED.
 CC EMBL: M84237; AAA74402.1; -
 CC EMBL: U33879; AAA79832.1; -
 CC EMBL: M34189; AAA59182.1; -
 CC EMBL: U33880; AAA79833.1; -
 CC EMBL: U33879; AAA79833.1; JOINED.
 CC EMBL: U28252; AAA81366.1; -
 CC EMBL: U33882; AAA79834.1; -
 CC EMBL: U33879; AAA79834.1; JOINED.
 CC EMBL: U33881; AAA79834.1; JOINED.
 CC EMBL: M84237; AAA74403.1; -
 CC PIR: B27079; B27079.
 CC HSP: P15358; 1SKZ.
 CC MW: 135630; -
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR002369; integrin_B.
 CC InterPro: IPR001169; integrin_beta.C.
 CC InterPro: IPR003659; PSI.
 CC InterPro: IPR002035; WFA.
 CC Pfam: PF00362; integrin_B.1.
 CC PRINTS: PR01186; INTEGRIN.
 CC PRODOM: PD001811; integrin_B.1.
 CC SMART: SM00001; EGF-like; 1.
 CC SMART: SM00187; INB; 1.
 CC SMART: SM00423; PSI; 1.

DR SMART: SM00327; VWA: 1.
 DR PROSITE; PS00243; INTEGRIN_BETA: 3.
 DR PROSITE; PS00222; EGF_1; UNKNOWN: 2.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; glycoprotein;
 KW Repeat; Signal; Phosphorylation; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 798 INTEGRIN BETA-1.
 FT DOMAIN 21 728 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 729 751 POTENTIAL.
 FT DOMAIN 752 798 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 140 378 WFMA-LIKE.
 FT DOMAIN 466 635 4 CYSTEINE-RICH TANDEM REPEATS.
 FT REPEAT 466 515 I.
 FT REPEAT 516 559 II.
 FT REPEAT 560 598 III.
 FT REPEAT 599 635 IV.
 FT DISULFID 27 464 BY SIMILARITY.
 FT DISULFID 35 45 BY SIMILARITY.
 FT DISULFID 38 75 BY SIMILARITY.
 FT DISULFID 48 64 BY SIMILARITY.
 FT DISULFID 207 213 BY SIMILARITY.
 FT DISULFID 261 301 BY SIMILARITY.
 FT DISULFID 401 415 BY SIMILARITY.
 FT DISULFID 435 466 BY SIMILARITY.
 FT DISULFID 462 466 BY SIMILARITY.
 FT DISULFID 477 489 BY SIMILARITY.
 FT DISULFID 486 525 BY SIMILARITY.
 FT DISULFID 491 500 BY SIMILARITY.
 FT DISULFID 502 516 BY SIMILARITY.
 FT DISULFID 531 536 BY SIMILARITY.
 FT DISULFID 533 568 BY SIMILARITY.
 FT DISULFID 538 553 BY SIMILARITY.
 FT DISULFID 555 560 BY SIMILARITY.
 FT DISULFID 574 579 BY SIMILARITY.
 FT DISULFID 576 607 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 FT DISULFID 592 599 BY SIMILARITY.
 FT DISULFID 613 618 BY SIMILARITY.

Query Match 44.1%; Score 1831; DB 1; Length 798;
 Best Local Similarity 44.9%; Pred. No. 1.2e-116;
 Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

QY 3 NICTRGVSSCOCLAVSPMCWCD---EALPLGSPRCDELKKNLKNCAPESEIFP 57
 DB NICTRGVSSCOCLAVSPMCWCD---EALPLGSPRCDELKKNLKNCAPESEIFP 57
 QY 25 NNCILANAKSCGCEICQAGPNCWCJNSFTLQEGMT-SARCDDEALKKKCGPPDINP 83
 DB NNCILANAKSCGCEICQAGPNCWCJNSFTLQEGMT-SARCDDEALKKKCGPPDINP 83
 QY 58 VSEARVLEDRPLSDKSGDSSQ---VTQVSPRIALRLRPDSKNFSIOVROVEDYPV 112
 DB VSEARVLEDRPLSDKSGDSSQ---VTQVSPRIALRLRPDSKNFSIOVROVEDYPV 112
 QY 84 RGSKDILKKKNVTNRSKGTAEKLPEDIHQIOPQOLVLRKSGEPQTFELKRAEDYPI 143
 DB RGSKDILKKKNVTNRSKGTAEKLPEDIHQIOPQOLVLRKSGEPQTFELKRAEDYPI 143
 QY 113 DIYYLMDLSYMKDLMSTQNLGTALQMKLTSLNLRIGFAPFQDKVVPYMYISPEEA 172
 DB DIYYLMDLSYMKDLMSTQNLGTALQMKLTSLNLRIGFAPFQDKVVPYMYISPEEA 172
 QY 144 DIYYLMDLSYMKDLMSTQNLGTALQMKLTSLNLRIGFAPFQDKVVPYMYISPEEA 202
 DB DIYYLMDLSYMKDLMSTQNLGTALQMKLTSLNLRIGFAPFQDKVVPYMYISPEEA 202
 QY 173 LNPCCVDMKTTCTLPMEGYKHVTLFDOVYTRFNEEKKOSVSRNRDAPGSGDAIMQATVC 232
 DB LNPCCVDMKTTCTLPMEGYKHVTLFDOVYTRFNEEKKOSVSRNRDAPGSGDAIMQATVC 232
 QY 203 LNPCCVDMKTTCTLPMEGYKHVTLFDOVYTRFNEEKKOSVSRNRDAPGSGDAIMQATVC 261
 DB LNPCCVDMKTTCTLPMEGYKHVTLFDOVYTRFNEEKKOSVSRNRDAPGSGDAIMQATVC 261
 QY 233 DERIGRNDASHLLVFTTDAKTHIALDGLAGIYVOPNDGCHGVSDNHYASTMDYPSL 292
 DB DERIGRNDASHLLVFTTDAKTHIALDGLAGIYVOPNDGCHGVSDNHYASTMDYPSL 292
 QY 262 GSLIGWRN-VTRLVSTFDAGHFAGDGKLGIVLPNDQCHL-ENNMVYTMHHYDYPSI 319
 DB GSLIGWRN-VTRLVSTFDAGHFAGDGKLGIVLPNDQCHL-ENNMVYTMHHYDYPSI 319
 QY 293 GLMTEKLSQKNINLFAVYENNVNLYXONSELIPCTGYVLSMDSNNVQLIYDAVYKIR 352
 DB GLMTEKLSQKNINLFAVYENNVNLYXONSELIPCTGYVLSMDSNNVQLIYDAVYKIR 352
 QY 320 AHLVYKLSNNIQTIFAVTEEPQVYKELKNLIPSAVGTSSANSNNVQLIYDAVYKIR 379
 DB AHLVYKLSNNIQTIFAVTEEPQVYKELKNLIPSAVGTSSANSNNVQLIYDAVYKIR 379
 QY 353 SKVTELEVRDLPELSTLSPNATCN--NEVYPLKSCMGKIDVYFSFSEAKVRCPOEK 410
 DB SKVTELEVRDLPELSTLSPNATCN--NEVYPLKSCMGKIDVYFSFSEAKVRCPOEK 410
 QY 380 SEVILENGKLSREVITISYCKNGVNGEGRKCSNLSIDVQFELSTISNKKCPKKD 439
 DB SEVILENGKLSREVITISYCKNGVNGEGRKCSNLSIDVQFELSTISNKKCPKKD 439
 QY 411 EKSFTIKPVGFKDSLIYQVTFDQDCAQQAQAPNSHRCNNGNGTEGCGVRCGPGMLSQ 470
 DB EKSFTIKPVGFKDSLIYQVTFDQDCAQQAQAPNSHRCNNGNGTEGCGVRCGPGMLSQ 470

DB 440 SDSKIRPLIGFTEVEVILLQYICECEGSEGIDESPKEHGNTEGACRONEGRVGRH 499
 QY 471 CESEEDYRPSQDE-CSPREGQPVCSQREGLCGGCVCHSSDPFK--ITGRYCEDDPS 527
 DB CESEEDYRPSQDE-CSPREGQPVCSQREGLCGGCVCHSSDPFK--ITGRYCEDDPS 527
 QY 500 CECSSTVEVSEDDAICRKNSEISNNGECGCGVCRKRDNREITSKGFCECDNFN 559
 DB CECSSTVEVSEDDAICRKNSEISNNGECGCGVCRKRDNREITSKGFCECDNFN 559
 QY 528 CVRRKEMSCSGHGCSCGDLCDSDMTGYCNCOTTTDTCMSSNGLICSGRKGCEGSCV 587
 DB CVRRKEMSCSGHGCSCGDLCDSDMTGYCNCOTTTDTCMSSNGLICSGRKGCEGSCV 587
 QY 560 CDPSNGLICGNVCKRCRCVCECPNNTGSAQDCLDSTCEASNGICNGRIGCEGVCK 619
 DB CDPSNGLICGNVCKRCRCVCECPNNTGSAQDCLDSTCEASNGICNGRIGCEGVCK 619
 QY 588 CLOPGSYGDTCEKCPDPACTPFKCEVCKEFCRGALEHNTCNRYCR----DELESVK 643
 DB CLOPGSYGDTCEKCPDPACTPFKCEVCKEFCRGALEHNTCNRYCR----DELESVK 643
 QY 620 CTDPKFGQGTCEMCQTCITGLVCAEHKECVQCARFNKGEKRD--TCTQECYSFNTYESKD 677
 DB CTDPKFGQGTCEMCQTCITGLVCAEHKECVQCARFNKGEKRD--TCTQECYSFNTYESKD 677
 QY 644 EL-KDQGXDAV-NCTYKNEDDQVRFROYEDSSGSIILVVEPEECRGPDIILVLSYV 701
 DB EL-KDQGXDAV-NCTYKNEDDQVRFROYEDSSGSIILVVEPEECRGPDIILVLSYV 701
 QY 678 KLPQVOPDPVSHCKEKEDDDCFYTTYSYNGNENVAHVVEPECPIDPIIPIVAGV 737
 DB KLPQVOPDPVSHCKEKEDDDCFYTTYSYNGNENVAHVVEPECPIDPIIPIVAGV 737
 QY 702 GAILLIGLALLIMKLLITTHDRKEPAKFEERARAKMDTANNPLYKATSTFTNITYRG 761
 DB GAILLIGLALLIMKLLITTHDRKEPAKFEERARAKMDTANNPLYKATSTFTNITYRG 761
 QY 738 AGVILGLALLIMKLLITTHDRKEPAKFEERARAKMDTANNPLYKATSTFTNITYRG 797
 DB AGVILGLALLIMKLLITTHDRKEPAKFEERARAKMDTANNPLYKATSTFTNITYRG 797

RESULT 11
 ITBO_XENLA STANDARD; PRT; 798 AA.
 ID ITBO_XENLA STANDARD; PRT; 798 AA.
 AC P12607;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-1* precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP MEDLINE=8816829; PubMed=2833505;
 RA Desimone D.W., Hynes R.O.;
 RT "Xenopus laevis integrins. Structural conservation and evolutionary
 divergence of integrin beta subunits";
 RL J. Biol. Chem. 263:5333-5340(1988).
 CC -1- FUNCTION: BETA INTEGRINS ASSOCIATE WITH ALPHA SUBUNITS TO FORM
 RECEPTOR COMPLEXES THAT RECOGNIZE THE SEQUENCE R-G-D IN A WIDE
 ARRAY OF LIGANDS (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- INTEGRIN BETA-1 AND BETA-1* OF XENOPUS ONLY DIFFER BY 19 AMINO
 ACIDS.
 CC -1- SIMILARITY: CONTAINS 1 WFMA-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M20180; AAA4890.1; -.
 DR PIR; B28193; B28193.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002369; Integrin_B.
 DR InterPro; IPR001168; Integrin_beta_C.
 DR InterPro; IPR003659; PSI.
 DR InterPro; IPR002035; WFMA.
 DR Pfam; PF00362; Integrin_B_1.
 DR PRINTS; PR01186; INTEGRINB.
 DR PRODOM; PD001811; Integrin_B_1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM00423; PSI; 1.

DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 DR InterPro: PS00022; EGF_1; UNKNOWN_2.
 KW InterPro: Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat: signal; Phosphorylation.
 FT SIGNAL 1 21
 FT CHAIN 22 798
 FT DOMAIN 22 727
 FT TRANSMEM 728 751
 FT DOMAIN 752 798
 FT DOMAIN 139 377
 FT DOMAIN 466 635
 FT REPEAT 466 635
 FT REPEAT 516 559
 FT REPEAT 516 559
 FT REPEAT 559 635
 FT REPEAT 599 635
 FT DISULFID 28 464
 FT DISULFID 36 46
 FT DISULFID 39 76
 FT DISULFID 49 65
 FT DISULFID 206 212
 FT DISULFID 260 300
 FT DISULFID 400 414
 FT DISULFID 434 601
 FT DISULFID 462 466
 FT DISULFID 477 489
 FT DISULFID 486 525
 FT DISULFID 491 500
 FT DISULFID 502 516
 FT DISULFID 531 536
 FT DISULFID 533 568
 FT DISULFID 538 553
 FT DISULFID 555 560
 FT DISULFID 574 579
 FT DISULFID 576 607
 FT DISULFID 581 590
 FT DISULFID 592 599
 FT DISULFID 613 618
 FT DISULFID 615 661
 FT DISULFID 620 630
 FT DISULFID 633 636
 FT DISULFID 640 649
 FT DISULFID 646 723
 FT DISULFID 665 699
 FT MOD_RES 783 783
 FT CARBOHYD 109 109
 FT CARBOHYD 131 131
 FT CARBOHYD 211 211
 FT CARBOHYD 223 223
 FT CARBOHYD 268 268
 FT CARBOHYD 362 362
 FT CARBOHYD 416 416
 FT CARBOHYD 481 481
 FT CARBOHYD 520 520
 FT CARBOHYD 584 584
 FT CARBOHYD 669 669
 SQ SEQUENCE 798 AA; 88303 MW; A1C45EA3711CF7C7 CRC64;
 Query Match 44.0%; Score 1826.5; DB 1; Length 798;
 Best Local Similarity 44.8%; Pred. No. 2.4e-116;
 Matches 350; Conservative 127; Mismatches 276; Indels 29; Gaps 15;

DB 144 LYLMDLSFSKMDLENNKSLGTLALMTMEKITSDFRIGSGEWEKTYMPISTTPAKLL 203
 QY 174 ENPCYDMKTTCLPMGEYKAVLTLDQYTRFNEEYKOSVSHNRDAPCEGDAINQATVCD 233
 DB 204 -NPCIN-DQNTSPSPSYKNVNLTRDKGLFNDLVGKQOISNLDSPGEGFADINQAVCG 261
 QY 234 EKIGRNDASHLLVFTTDAKTHIALDGLAGIYVDPDQCHVSSDNHYSASTYMDYPSLG 293
 DB 262 E0IGRNR-VTRLVFSTAGFHAGDGLGIVPNDGRCHL-HGNMYTMSHYDPSIA 319
 QY 294 LMTESLQKNLIFAVENYVNLQYNSSELIPGTGVLSMUSSNVQLIVADYAGKRS 353
 DB 320 HLVQKLSNNIOTITPAVEDPOPIYQELKNLIPASAGTSSNSNVYQLIIDSLS 379
 QY 354 KVELEVRDLPEELSLFNATCLNNEVTPGL---KSCMGLKIGVTSLSLEAKYRGCPQE 409
 DB 380 ELLIENSLKPEGVITSYRSPCKNG--VKGTEGDEKKSNIIS1GQVREISVFAHKCPK 437
 QY 410 -KEKSFTRPVGFKDSLIVQYTFDCCACQAOAEPNSHRKNGSTTECCYCRGCPWLG 468
 DB 438 GQAESIKIKPLGFNEVEIVLQFCECDODKGTTPNSPECHFGNGTECGACRONEGRIG 497
 QY 469 SQCESEDDYRPSQODE--CSPREGOPVCSQREGCLGCGVCCHSSDPFK--ITGRVCECD 525
 DB 498 KECBSTDEVNSSEDMDAYCRRENSSEISNNGDCICGCVCKKNDPNREYVSGATCECDN 557
 QY 526 FSCVYKKEEMSGHSGCSCDCLSDMTGYCNCCTTPTDCHSNGLLSGRKEGCS 585
 DB 558 FNCDSNLLIGGKGVCKGCRVCECPNYSGSACDSSEDTSCMAKNGICNGRIDCGR 617
 QY 586 CYCLOPGSYGDTCEKCPGPCACTFKKCEVCKKFKDGLDENTCNRYCR---DELES 641
 DB 618 CKCTPCKPGCPCELCQTCVGVCTEKKECVOCRAFOKEKOD--VCMEQCHMFINSLVDS 675
 QY 642 VKELDTG--KDAVCTKKNEDDCVFRQYEDSSGKILVVEPECPKPDILVLLS 699
 DB 676 REELQPGQAEALYTKCKEKDAEDCFYTYVSDSNEMVAVHYEPREPSPDIIPIVAG 735
 DB 736 VVAGIVLIGLALLIMKLLIMLIIHRRERAFKEKEMAKMDTGEMPIYKSAVAVNPKY 795
 QY 760 RG 761
 DB 796 EG 797
 RESULT 12
 ID ITB1_MOUSE STANDARD; PRT; 798 AA.
 AC P09055;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
 DE (CD29) (Integrin VLA-4 beta subunit).
 GN ITGB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=89005707; Pubmed=3262537;
 RA Tomihata S;
 RT "Murine mRNA for the beta-subunit of Integrin is increased in
 BAIB/c-3T3 cells entering the G1 phase from the G0 state."
 RL FEBS Lett. 238:315-319(1988).
 RN [2]
 RP SEQUENCE OF 2-798 FROM N.A.
 RC STRAIN=BAIB/C;
 RX MEDLINE=89235580; Pubmed=2523953;

Query Match	Best Local Similarity	Score	DB 1;	Length	798;
Matches	346;	Conservative	137;	Mismatches	269;
				Indels	28;
				Gaps	
QY 3	NICTRGRAYSSCOCCIAVSPMCANCS	43.88;	Score 1821;	DB 1;	Length 798;
Db 25	NKCLIKANAKSGEGCIQAPNGMCTNTTFLDEGMPT-SARCDDEALKKKGCOPSDIENP	44.48;	Pred. No. 5,7e-116;		
QY 58	VSEARVLDRPLSPDGSDSSQ-----VTQVSPORILRLRPDDSKNFSIOVROVEDPY	44.48;	Pred. No. 5,7e-116;		
Db 84	RGSQITKKNNKVNTRSKMAKRLRPEDITQIQPOULLIKLRSGPQKFTLKFRAEDYPI	44.48;	Pred. No. 5,7e-116;		
QY 113	DIYLLMDLSYMKDDLSIOMLGTKLATQMRKLTSLNLRIGFAGVDRPVSPYMIISPEEA	44.48;	Pred. No. 5,7e-116;		
Db 144	DLYYLLMDLSYMKDDLENVRSLGIDLMENMRITSDFRIGSGFEVETKMPYISTPBAK-2020	44.48;	Pred. No. 5,7e-116;		
QY 173	LENPCYDMKTTCLPMFEGYKHVLTITDOYTRNENYKKQVSRRNDAPDEGFDALMOATVC	44.48;	Pred. No. 5,7e-116;		
Db 203	LRNPC-TSEQNCISPFYSKYNLSTLDKREFENELVGOQRISGNLSDSPREGFDAIMOVAVC	44.48;	Pred. No. 5,7e-116;		
QY 233	DEKIGMRNDASHLVFTTDAKTHALDGRLAGIYQPPNDGCGHVSNDHYASASTMYPSL	44.48;	Pred. No. 5,7e-116;		
Db 262	GSLLGMRN-VTRLVFTSDAGFHRAGQKLGGLVLPNDGCHL-ENNVTYMSHYDPST	44.48;	Pred. No. 5,7e-116;		
QY 293	GIAMEKISQKINILIFATTEWVWVLXONYSELIPGTIVGLSMDSSVNLQIYDAVGKIR	44.48;	Pred. No. 5,7e-116;		


```

FT DISULFID 412 668 BY SIMILARITY.
FT DISULFID 439 443 BY SIMILARITY.
FT DISULFID 454 466 BY SIMILARITY.
FT DISULFID 463 502 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 478 493 BY SIMILARITY.
FT DISULFID 508 513 BY SIMILARITY.
FT DISULFID 510 545 BY SIMILARITY.
FT DISULFID 515 530 BY SIMILARITY.
FT DISULFID 532 537 BY SIMILARITY.
FT DISULFID 551 556 BY SIMILARITY.
FT DISULFID 553 584 BY SIMILARITY.
FT DISULFID 558 567 BY SIMILARITY.
FT DISULFID 569 576 BY SIMILARITY.
FT DISULFID 590 595 BY SIMILARITY.
FT DISULFID 592 638 BY SIMILARITY.
FT DISULFID 597 607 BY SIMILARITY.
FT DISULFID 610 613 BY SIMILARITY.
FT DISULFID 617 626 BY SIMILARITY.
FT DISULFID 623 700 BY SIMILARITY.
FT DISULFID 642 676 BY SIMILARITY.
FT MOD.RES 760 760 PHOSPHORYLATION (BY SIMILARITY).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 773 773
SQ SEQUENCE 773 AA; 85322 MM; 54340886CE157195 CRC64;

```

Query Match 43.7%; Score 1817; DB 1; Length 773;
 Best Local Similarity 44.2%; Pred. No. 1e-115;
 Matches 344; Conservative 138; Mismatches 268; Indels 28; Gaps 13;

```

QY 3 NICTTREVSSCOCLAVSPMCAMCSD-----EALPLGSPDCDLKENLKNCAPESEIEP 57
DB 2 NRCIKANAKSCGECIQAGPCMGCTNSTFLQEGMPT-SARCDLEALKKKCHPNDIENP 60
QY 58 VSEARVEDRPLSPKSGSSGSSO-----VQVSPQRIALRLRPDSKFSIQVROVEDYV 112
DB 61 RGSFOIKRKNKVNTRSGTAETKLOPEDITQIQPOQLVLRLSGEPQFTFLKRAEDYPT 120
QY 113 DLYVLMDSYMKDDLSIONLGTKLATQMRKLTSLNLIGFADFVDPSPVYIISPPEA 172
DB 121 DLYVLMDSYMKDDLSIONLGTKLATQMRKLTSLNLIGFADFVDPSPVYIISPPEA 172
QY 173 LNPCTYMKTTCLPMEGYKAVLTLDQVYTRPNEVKKQSSRNADAEFGFDALIMQATVC 232
DB 180 LNPCTYMKTTCLPMEGYKAVLTLDQVYTRPNEVKKQSSRNADAEFGFDALIMQATVC 232
QY 233 DEKGMWNRDASHLVTFTDAKTHALRAGIYQPNDDGCHGSDNHYASTMDVPSL 292
DB 239 GSLIGMWN-VTRLVSTDAVHFAGAGKGLGIVLPDGGCHLEND-VYMSHYIDYPT 296
QY 293 GLMTEKLSOKNINLFAVTENVNLYONYSELJPGTTGVLSMDSNVNLIIDVAYGKIR 352
DB 297 AHLVOKLSENNIQTIFAVTEEFQVYKELKILPKSAVGTLSANNSVNIQIIDIAYNSLS 356
QY 353 SKVELEPRDLPBELSLSFNATCLN-NEVLPGLKSCMGKLTGQVVSISIAKRGCPQEK 410
DB 357 SEVILENSKLAEGVTINYSKYCKNGVNGTEGNGKCSNISIGDVOEISITANKCPKNK 416
QY 411 EKSTFIPVGEKSLIVQVFPDCACAOAQAEPHSRRCNNGNGTFFEGVRCGCGWLSG 470
DB 417 SETIKIRPLGTEVEVEIITQICBCEGCEITPOSPPCHGNGTFFEGVRCGCGWLSG 476

```

```

QY 471 CECSEEDYRPSQODE-CSPREGQVYCSORGELCGQCVCHSSDFGK--ITGKYCECDFES 527
DB 477 CECSTDEVNSEMDYAYCKRNSSEICSNNGCEVCQGCRCRRDNTNELYSKPECECNFN 536
QY 528 CVRYKGEKSCSHGCGSCGDDCLDSDMTGYVNCNTRDTQMSNGLLCSGKCEGCSGV 587
DB 537 CDRSNGLLCGGNGVCVCEPCNPNYTSACDCSLGTSMAVNGQICNGVCECGACK 596
QY 588 CIOPEYSYDPECEKPTCPDCTFFKCEVECKKFDGALHDENTGNRYCR---DEISVK 643
DB 597 CDDPFGPTCEMCTCGVCAEHKECVQCAFPNNGEKKD--TCAQECSHLNTIKVENRD 654
QY 644 ELKDTGK--DAVNCTYKKNEDPCVNRQYTEDSSGKSLIYVVEEPCPGPDIIVLSVM 701
DB 655 KLPQGGQVPLSHCKEKEDVDCEMFEFTYSVNGNNATVAVVETCEPDIIPVAGVV 714
QY 702 GALLIGLALLIWLTLTTHDRKEFAFEERARAKDTPNPNLYKATSTFNITY 759
DB 715 AGIVLIGLALLIWLTLTTHDRKEFAFEERARAKDTPNPNLYKATSTFNITY 772

```

```

RESULT 15
ID ITBI_RAT STANDARD: PRT: 799 AA.
AC P49134.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-1 precursor (Fibronectin receptor beta subunit)
DE (CD29) (Integrin VLA-4 beta subunit).
GN ITGB1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95331632; PubMed=7541764;
RA Malek-Hedayat S., Rome L.H.;
RT Cloning and sequence of the cDNA encoding the rat oligodendrocyte
RT integrin beta 1 subunit."
RL Gene 158:287-290(1995)
CC -1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-
CC 1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-
CC 1/BETA-1 AND ALPHA-2/BETA-1 RECOGNIZE THE PROLINE-HYDROXYLATED
CC SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-
CC 3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-
CC 10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR
CC FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN
CC THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN.
CC INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN
CC ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1
CC ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR
CC FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-I-D-S IN VCAM1. INTEGRIN
CC ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOTOXIN AND
CC OSTEOPOINTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-C-I-E-L IN
CC CYTOTOXIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPLIGRIN
CC AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR
CC VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A
CC WIDE ARRAY OF LIGANDS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1
CC ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-
CC 5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR
CC ALPHA-V.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U12309; AAA86669.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PF00362; Integrin_B.1.
 DR PRINTS: PR01186; INTEGRIN_B.1.
 DR ProDom: PD001811; Integrin_B.1.
 DR SMART: SM00001; EGF-like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal; Phosphorylation.
 FT SIGNAL 1 20
 FT CHAIN 21 798
 FT DOMAIN 21 729
 FT TRANSMEM 730 752
 FT DOMAIN 753 798
 FT DOMAIN 140 378
 FT DOMAIN 467 636
 FT REPEAT 467 516
 FT REPEAT 517 560
 FT REPEAT 561 599
 FT REPEAT 600 636
 FT DISULFID 27 465
 FT DISULFID 35 45
 FT DISULFID 38 75
 FT DISULFID 48 64
 FT DISULFID 207 213
 FT DISULFID 261 301
 FT DISULFID 401 415
 FT DISULFID 435 692
 FT DISULFID 463 490
 FT DISULFID 478 490
 FT DISULFID 487 526
 FT DISULFID 492 501
 FT DISULFID 503 517
 FT DISULFID 532 537
 FT DISULFID 534 569
 FT DISULFID 539 554
 FT DISULFID 556 561
 FT DISULFID 575 580
 FT DISULFID 577 608
 FT DISULFID 582 591
 FT DISULFID 593 600
 FT DISULFID 614 619
 FT DISULFID 616 662
 FT DISULFID 621 631
 FT DISULFID 634 637
 FT DISULFID 641 650
 FT DISULFID 647 724
 FT DISULFID 666 700
 FT MOD. RES 784 784
 FT CARBOHYD 50 50
 FT CARBOHYD 94 94
 FT CARBOHYD 97 97
 FT CARBOHYD 212 212
 FT CARBOHYD 269 269
 FT CARBOHYD 363 363
 FT CARBOHYD 406 406
 FT CARBOHYD 417 417
 FT CARBOHYD 482 482
 FT CARBOHYD 521 521
 FT CARBOHYD 585 585
 FT CARBOHYD 670 670
 FT DISULFID 27 465 BY SIMILARITY.
 FT DISULFID 35 45 BY SIMILARITY.
 FT DISULFID 38 75 BY SIMILARITY.
 FT DISULFID 48 64 BY SIMILARITY.
 FT DISULFID 207 213 BY SIMILARITY.
 FT DISULFID 261 301 BY SIMILARITY.
 FT DISULFID 401 415 BY SIMILARITY.
 FT DISULFID 435 692 BY SIMILARITY.
 FT DISULFID 463 490 BY SIMILARITY.
 FT DISULFID 478 490 BY SIMILARITY.
 FT DISULFID 487 526 BY SIMILARITY.
 FT DISULFID 492 501 BY SIMILARITY.
 FT DISULFID 503 517 BY SIMILARITY.
 FT DISULFID 532 537 BY SIMILARITY.
 FT DISULFID 534 569 BY SIMILARITY.
 FT DISULFID 539 554 BY SIMILARITY.
 FT DISULFID 556 561 BY SIMILARITY.
 FT DISULFID 575 580 BY SIMILARITY.
 FT DISULFID 577 608 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT DISULFID 593 600 BY SIMILARITY.
 FT DISULFID 614 619 BY SIMILARITY.
 FT DISULFID 616 662 BY SIMILARITY.
 FT DISULFID 621 631 BY SIMILARITY.
 FT DISULFID 634 637 BY SIMILARITY.
 FT DISULFID 641 650 BY SIMILARITY.
 FT DISULFID 647 724 BY SIMILARITY.
 FT DISULFID 666 700 BY SIMILARITY.
 FT MOD. RES 784 784 PHOSPHORYLATION (BY SIMILARITY).
 FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 521 521 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 585 585 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 799 AA; 88494 MW; F4475202EB8A3EA6 CRC64;

Query Match 43.4%; Score 1801.5; DB 1; Length 799;

Best Local Similarity 44.2%; Pred. No. 1.2e-114;

Matches 345; Conservative 136; Mismatches 271; Indels 29; Gaps 14;

3 NICTTGAVSSCOOCLAVSPMCAMCSD----FALPLGSPROCLKENLKDNCAPSIIEFP 57
 25 NRCLKRNKAKSCGCECLOAGNCGMCTTTFLQEGMPT-SARCDLALKKKGGHPSDLENP 83
 58 VSEARVLEDRLPLSDKSGDSO-----VQVSPRIARLRPDSKNFSIOGVQVEDPY 112
 84 RSGQITKRNKNTVNSKGAKEKLRPEDITLIQIPQOLLKLRGEPKFTLAKFRADYPI 143
 113 DLYYLMDSYSKKDLMSLQNGTALQAMRKLTSLNLRGAFYDKVSPRMTISPEEA 172
 144 DLYYLMDSYSKMDLENKSLGTDLMEMRRTSDFRIGFSVEKTYMPTISTPAK- 202
 173 LBNPCYDMKTTLPLMGYKHYVLTLDQVTRFNEEYKKSQVSRNRPAPGCFDAIMQAVTC 232
 203 LNPCC-TSPQNTCTSPFYKKNVSLTDRGFEHLYGOORISGNDSPRGCFDAIMQAVTC 261
 233 DEKIGRNDASHLVLFTTDAKTHIALDGRLAGIYOPNDGQCHVSGDNNHYSASTMDYPSL 292
 262 GSLIGMRN-VTRLVFSTDFAGFHAGDGKLGIVLPNDGCHL-ENNVYTMSHYVDYPSI 319
 293 GLMTEKLSQKNLNLFEATENVVNLXONSLEICTYGVLSMSSNVLOLIVAYGIR 352
 320 AHLVOKLSNNNTQTLFEATVEEFOYKELKNLPISAVGTLGSSNVYLOLIDAYSL 379
 353 SKVELEVRDLPEBLSLSNATCLN--NVEYPLGKSCMLKIDYVFSIEKVGCP-QE 409
 380 SEVILLENKLPDGVITINYSCKKNGVNGTENGKRCNSISIGDEVQFISTANKPKKE 439
 410 KEKSTTIKPVGKSDLIYQVTFDCCACQAAEPNSHPCNNNGNGTEECVCRGPGMLGS 469
 440 SENQKLNPDLGTFTEVEVYVLOFICNCSQSIGIPASPCCHGNGTFECGACRCEGRGR 499
 470 QCESEEDYRPSQOQE-CSPREGPVCSORGECLOGCVCVCHSSPDGK--ITGKYCECDNF 526
 500 HECCTDEVNSMDMAYCRKKESSITCSNNECVCQGCYCKKRENTNITSYKCECDNF 559
 527 SCVRYKGMCSGHCSCGCDLCSDDWTGYVCNCTRTTDTMCSSNGLCSGRKCEGSC 586
 550 NCDRNGNLICGNGVCRGRCVCEYPNTGSACDSLDTPCVATNGOICNGRGICEGAC 619
 587 VCIQPGSYGDPCCKPTCPDACTPKKECEVCKKEDRGLHDEMNCKRYC---DEIESV 642
 620 KCTDKREGPTCECTQTLGVCAEHKECVOCRAFNKEGRKD--TCAQECSHFNLTKEYSR 677
 643 KEIKDTGR--DAVNCTYENEDCVVRFOYEDSSKSTLYVEPECPKGDIVLVLSV 700
 678 EKLPPQVADVDTYHCKEKDIDDCAFYFYSVNSGGEAVHVEYPPDCEPTGDIPIYAGV 737
 701 MGAILLGLAALLIMKLLITTHDRKEFAKEEERAKAKMDTANNPLYLEATSTFTNTYR 760
 738 VAGIYVIGLALLIMKLLIMTHDRREFAKFEKEMNKKMDGEMPIYKSAVTTVVNPXE 797
 761 G 761
 798 G 798

Search completed: May 19, 2002, 12:27:36
 Job time: 6470 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 10:38:58 ; Search time 96.49 Seconds
(without alignments)
1366.174 Million cell updates/sec

Title: US-09-673-302A-1
4154
Sequence: 1 GPNICTRGVSSCQCLAVS.....NNPLYKEATSTFNITVYRGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4016	96.7	784	6	097702
2	4013	96.6	784	6	097702
3	3996	96.2	788	6	097703
4	3898	93.8	784	6	0957H1
5	3894	93.7	784	6	097705
6	3606	86.8	723	11	096W16
7	3484.5	83.9	781	11	092071
8	3472	83.6	680	11	096W15
9	3260.5	78.5	788	13	007012
10	2315.5	55.7	791	6	096K49
11	1833	44.1	798	6	096GP0
12	1800.5	43.3	807	13	097A01
13	1658	39.9	788	5	096444
14	1609	38.7	837	5	09N4S7
15	1588.5	38.2	783	5	092163
16	1584	38.1	799	5	095P95

17	1553	37.4	772	13	092070
18	1508	36.3	806	5	061677
19	1506	36.3	772	13	097PU4
20	1505	36.2	806	5	091808
21	1498	36.1	764	6	097U04
22	1498	36.1	802	5	076727
23	1420.5	34.8	767	13	0987H6
24	1403.5	33.8	768	13	0987H8
25	1390.5	33.3	792	5	017494
26	1382	33.5	801	5	091774
27	1312	31.6	790	5	09GSR3
28	1225.5	29.5	640	11	009182
29	1145	27.6	878	5	097189
30	1080.5	26.0	738	11	088424
31	1079.5	26.0	838	5	018482
32	1064	25.6	764	5	097343
33	828.5	19.9	439	4	09BUC9
34	754.5	18.2	290	11	09D214
35	738	17.8	150	11	063846
36	603	14.5	134	6	096M51
37	419	10.1	194	13	091415
38	417.5	10.1	403	4	014549
39	417.5	10.1	494	4	095965
40	368	8.9	69	6	0957C4
41	347	8.4	103	11	097483
42	327.5	7.9	166	6	018961
43	318	7.7	65	4	012806
44	304	7.3	114	11	09R151
45	281	6.8	94	6	095N85

ALIGNMENTS

RESULT 1
097702 PRELIMINARY: PRT: 784 AA.
AC 097702: 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DR 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PLATELET GLYCOPROTEIN IIA.
GN GPIIb.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99410043; PubMed=10482317;
RA Lipcomb D.L., Bourne C., Boudreau M.K.;
RT "DNA sequence of the canine platelet beta3 gene from cDNA: comparison
RT of canine and mouse beta3 to segments that encode alloantigenic sites
RT and functional domains of beta3 in human beings.";
RL J. Lab. Clin. Med. 134:313-321(1999).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC EMBL: AF116270; AAD13680.1; .
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_B.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR002035; VWFA.
DR Pfam: PF00362; Integrin_B.1.
DR PRINTS: PR01186; INTEGRINB.
DR PRODOM: PR001811; Integrin_B.1.
DR SMART: SM00001; EGF_Like; I.
DR SMART: SM00187; INB; I.
DR SMART: SM00423; PSI; I.
DR SMART; SM00327; VWA; I.

DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 KM Cell adhesion: Cytoskeleton: EGF-like domain; Extracellular matrix;
 KM Glycoprotein: Integrin: Repeat; Transmembrane.
 SQ SEQUENCE 784 AA; 8638 MW; 6B2A6D34916EA260 CRC64;

Query Match 96.7%; Score 4016; DB 6; Length 784;
 Best Local Similarity 96.3%; Pred. No. 0;
 Matches 731; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 3 NICTRGVSSCQCCLAIVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEFPVSEAR 62
 DB 25 NICTRGVHSCQCCLAIVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEFPVSEAR 84
 QY 63 VLEDRPLSDKSGSDSSQVTVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 122
 DB 85 ILLEAPLPSNKGSDSSQITQVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 144
 QY 123 SMKDDLSIONLGTKLATQMRKLTSLNLRIGGAFVDPKPSVYVISPPEALENFCYDMKT 182
 DB 145 SMDDLSIONLGTKLASQMHKLTSLNLRIGGAFVDPKPSVYVISPPEALENFCYDMKT 204
 QY 183 TCLPMFGYKHVLTLDQVTRFNEEVKQKOSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 242
 DB 205 TCLPMFGYKHVLTLDQVTRFNEEVKQKOSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 264
 QY 243 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHYSASTTMDYPSLGLMTEKLSQK 302
 DB 265 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHYSASTTMDYPSLGLMTEKLSQK 324
 QY 303 NINLIFAVTENVVNLQYNSSELPGTGVGLVSDSSNVQLIYDAVGKIRSKVELEVRDL 362
 DB 325 NINLIFAVTENVVNLQYNSSELPGTGVGLVSDSSNVQLIYDAVGKIRSKVELEVRDL 384
 QY 363 PEEILSLSFNATCLNNEVYIPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 422
 DB 385 PEEILSLSFNATCLNNEVYIPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 444
 QY 423 DSLIYQVTFDCCACQAOAEPNSHRCNNNGNTEFCGVCRCGPMGLSQCECSEEDYRPSQ 482
 DB 445 DSLIYQVTFDCCACQAOAEPNSHRCNNNGNTEFCGVCRCGPMGLSQCECSEEDYRPSQ 504
 QY 483 QDECSPREGQPVCSQREGLCGCCVCHSSDPFKITGKYCEDDFSCVARKGEMCSGHGQC 542
 DB 505 QDECSPREGQPVCSQREGLCGCCVCHSSDPFKITGKYCEDDFSCVARKGEMCSGHGQC 564
 QY 543 SCGDCLCDSMDWGYNCNCTTRDTFCMSSNGLLCGRGKCEGSCVCIOPGSYGDTECKCP 602
 DB 565 SCGDCLCDSMDWGYNCNCTTRDTFCMSSNGLLCGRGKCEGSCVCIOPGSYGDTECKCP 624
 QY 603 TCGDACTFKKECECKECPRGALHDENTNCRKCRDIESEYKELDTGKAVNCTYKNEDD 662
 DB 625 TCGDACTFKKECECKECPRGALHDENTNCRKCRDIESEYKELDTGKAVNCTYKNEDD 684
 QY 663 CVVRFOYEDSSGKSLIYVEEPECKGPDILVLLSVGAILLLGLALLIMKLLITTH 722
 DB 685 CVVRFOYEDSSGKSLIYVEEPECKGPDILVLLSVGAILLLGLALLIMKLLITTH 744
 QY 723 DRKEFAKFEEDERAKMDTANNDLYKEATSTFTNITYRGT 762
 DB 745 DRKEFAKFEEDERAKMDTANNDLYKEATSTFTNITYRGT 784

RESULT 2
 Q9TUN7 PRELIMINARY; PRT; 784 AA.
 AC Q9TUN7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 13, Last annotation update)

DE GLYCOPROTEIN GPIIIA.
 GN GPIIIA.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA Tao J., Parrilla R.;
 RT "Molecular cloning of dog GPIIIA cDNA."
 RL Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL, AF170525; AADA9737.1;
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR002035; vWFA.
 DR Pfam: PF00362; Integrin_B; 1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PST; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 KM Cell adhesion: Cytoskeleton: EGF-like domain; Extracellular matrix;
 KM Glycoprotein: Integrin: Repeat; Transmembrane.
 SQ SEQUENCE 784 AA; 86416 MW; 9D507F827628790A CRC64;

Query Match 96.6%; Score 4013; DB 6; Length 784;
 Best Local Similarity 96.2%; Pred. No. 0;
 Matches 731; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

QY 3 NICTRGVSSCQCCLAIVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEFPVSEAR 62
 DB 25 NICTRGVHSCQCCLAIVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESIEFPVSEAR 84
 QY 63 VLEDRPLSDKSGSDSSQVTVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 122
 DB 85 ILLEAPLPSNKGSDSSQITQVSPQIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDLSY 144
 QY 123 SMKDDLSIONLGTKLATQMRKLTSLNLRIGGAFVDPKPSVYVISPPEALENFCYDMKT 182
 DB 145 SMDDLSIONLGTKLASQMHKLTSLNLRIGGAFVDPKPSVYVISPPEALENFCYDMKT 204
 QY 183 TCLPMFGYKHVLTLDQVTRFNEEVKQKOSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 242
 DB 205 TCLPMFGYKHVLTLDQVTRFNEEVKQKOSVSRNRDAPEGGFDAIMQATVDEKIGMRNDA 264
 QY 243 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHYSASTTMDYPSLGLMTEKLSQK 302
 DB 265 SHLVFTTDAKTHIALDGRLAGIVQPDGQCHVSDNHYSASTTMDYPSLGLMTEKLSQK 324
 QY 303 NINLIFAVTENVVNLQYNSSELPGTGVGLVSDSSNVQLIYDAVGKIRSKVELEVRDL 362
 DB 325 NINLIFAVTENVVNLQYNSSELPGTGVGLVSDSSNVQLIYDAVGKIRSKVELEVRDL 384
 QY 363 PEEILSLSFNATCLNNEVYIPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 422
 DB 385 PEEILSLSFNATCLNNEVYIPGLKSCMGKLTIGDIVSFSIEAKVRGCPQEKESFTIKPVGFR 444
 QY 423 DSLIYQVTFDCCACQAOAEPNSHRCNNNGNTEFCGVCRCGPMGLSQCECSEEDYRPSQ 482

```

Db 445 DSIITVTTDCCACAOAEPSSHRGNNGTFEGCVCLCGPMTLGSQCECESEEDYHPSQ 504
Qy 483 QDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHQC 542
Db 505 QDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHQC 564
Qy 543 SCGDCLCDSMDWGYNCNCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEKCP 602
Db 565 SCGDCLCDSMDWGYNCNCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEKCP 624
Qy 603 TCPDACTFKECKEYCECKEFGALHDENTCNRYCRDIESEVKELKDTGDAVNCYKNEED 662
Db 625 TCPDACTFKECKEYCECKEFGALHDENTCNRYCRDIESEVKELKDTGDAVNCYKNEED 684
Qy 663 CVVRFQYEDSSGKSTLYVEEPECKPGDILVLLSVGAILLIGLALLIMKLITLTH 722
Db 685 CVVRFQYEDSSGKSTLYVEEPECKPGDILVLLSVGAILLIGLALLIMKLITLTH 744
Qy 723 DRKEFAKFEERARAKWDTANNPPLYKEATSTFTNTTYRGT 762
Db 745 DRKEFAKFEERARAKWDTANNPPLYKEATSTFTNTTYRGT 784

RESULT 3
O9TUN3 PRELIMINARY: PRT: 788 AA.
AC O9TUN3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN IIIA.
GN GPIIa.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLATELET;
RA Tao J., Parrilla R.;
RT "Molecular cloning of rabbit platelet GPIIa cDNA.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR EMBL: AF170529; AAD51955.1; -.
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; PSI.
DR Pfam: PR00362; Integrin_B; 1.
DR PRINTS: PRO1186; INTEGRINB.
DR PRODOM: PD001811; Integrin_B; 1.
DR SMART: SM00187; INB; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00243; INTEGRIN_BETA; 3.
KW Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix;
KW Glycoprotein; Integrin; Repeat; Transmembrane.
SQ SEQUENCE 788 AA; 87049 MW; 7ADB9CDA7301D78D CRC64;

```

```

Db 27 GPNICTTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCLKENLKDNCAPESIEFPVSE 86
Qy 61 ARVLEDRPLSKSGSGSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDL 120
Db 87 AOILEARPLSKSGSGSSQVTVSPORIALRLRPDDSKNFSIQVQVEDYPVDIYYLMDL 146
Qy 121 SYSMKDWLSIQNLGTYLKIATOMRKLTSNLRIGFAGYDKPVSPYMTSPPEALENFCYDM 180
Db 147 SYSMKDWLSIQNLGTYLKIATOMRKLTSNLRIGFAGYDKPVSPYMTSPPEALENFCYDM 206
Qy 181 KTCLPMFGKHYVLTLDQVTRFNEEYVKOSVSNRBPAPEGGFPAIMQATYCODEKIGWRN 240
Db 207 KTCLPMFGKHYVLTLDQVTRFNEEYVKOSVSNRBPAPEGGFPAIMQATYCODEKIGWRN 266
Qy 241 DASHLVFTTDAKTHIALDGRAGIIVQNDQCHVSGSDNHYSASTMDYPSLGMTKLS 300
Db 267 DASHLVFTTDAKTHIALDGRAGIIVQNDQCHVSGSDNHYSASTMDYPSLGMTKLS 326
Qy 301 QKNINLIFAVTEENVVNLQNTSELIPGTYGVLSMDSNVLLQILDVAGTKRSVLEVR 360
Db 327 QKNINLIFAVTEENVVNLQNTSELIPGTYGVLSMDSNVLLQILDVAGTKRSVLEVR 386
Qy 361 DLPEELISFNATCLNNEVITPGLKSCMGKIGDVPVSFIEAKVGRGPOEKEKSPFTIPVG 420
Db 387 DLPEELISFNATCLNNEVITPGLKSCMGKIGDVPVSFIEAKVGRGPOEKEKSPFTIPVG 446
Qy 421 FKDSLIVQVTFDCCACAOAEPNSHRCNNGNFTFEGCVCRGCGMTLGSQCESEEDYRP 480
Db 447 FKDSLIVQVTFDCCACAOAEPNSHRCNNGNFTFEGCVCRGCGMTLGSQCESEEDYRP 506
Qy 481 SQDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHG 540
Db 507 SQDECSFREGQPVCSQREGELCGQCVCHSSDFGKITGKYCECDDFSCVRYKGMCSGHG 566
Qy 541 QCSGDCCLCDSMDWGYNCNCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEK 600
Db 567 QCSGDCCLCDSMDWGYNCNCTTRDTGCMSSNGLLCSGRKCECCSCVCIOPGSYGDTCEK 626
Qy 601 CPTCPDACTFKECKEYCECKEFGALHDENTCNRYCRDIESEVKELKDTGDAVNCYKNE 660
Db 627 CPTCPDACTFKECKEYCECKEFGALHDENTCNRYCRDIESEVKELKDTGDAVNCYKNE 686
Qy 661 DDCVRFQYEDSSGKSTLYVEEPECKPGDILVLLSVGAILLIGLALLIMKLIT 720
Db 687 DDCVRFQYEDSSGKSTLYVEEPECKPGDILVLLSVGAILLIGLALLIMKLIT 746
Qy 721 IHDKEFAKFEERARAKWDTANNPPLYKEATSTFTNTTYRGT 762
Db 747 IHDKEFAKFEERARAKWDTANNPPLYKEATSTFTNTTYRGT 788

RESULT 4
O95JH1 PRELIMINARY: PRT: 784 AA.
AC O95JH1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN GPIIa.
GN CD61.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxId=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADIPOTIC SMOOTH MUSCLE;
RA Jimenez-Marin A.M., Garrido J.J., Llanes D., Barhanchio M.J.;
RT "Characterization of the porcine CD61 (GPIIa) gene.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF282890; AAK69529.1; -.
SQ SEQUENCE 784 AA; 86365 MW; 1AE756651EEF33B8 CRC64;

```

Query Match 93.8%; Score 3898; DB 6; Length 784;
 Best Local Similarity 92.8%; Pred. No. 0;
 Matches 707; Conservative 35; Mismatches 20; Indels 0; Gaps 0;

QY 1 GPNICITRGVSSCCOCLAVSPMCAMCSDDEALPLGSPRCDLKENILKNCAPESIEFPVSE 60
 DB 23 GPNICATRGVSSCCOCLAVSPMCAMCSDDEALPLGTPRCNLEENLNHCNPKIEFPISE 82
 QY 61 ARVLEDRPLSDKSGSDSSQTVSPQRIALRLPDDSKNFSIOVQVEDYPVDIYYLMDL 120
 DB 83 ARILEARPLSDKSGSDSSQTVSPQRIALRLPDDSKNFSIOVQVEDYPVDIYYLMDL 142
 QY 121 SYSKMDLMSIONLGTATLQATOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENFCYDM 160
 DB 143 SFSKMDLESIONLGTATLQATOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENFCYDM 202
 QY 181 KTTCLPMFGYKHHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRN 240
 DB 203 KTTCLPMFGYKHHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRN 262
 QY 241 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHYSASTTMDYPSLGLMTEKLS 300
 DB 263 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHYSASTTMDYPSLGLMTEKLS 322
 QY 301 QKNINLIFAVTVENYVNLQVNSSELIPGTGVLSMDSNVQLIYDAVYGRIRSKVELEVR 360
 DB 323 QKNINLIFAVTVENYVNLQVNSSELIPGTGVLSMDSNVQLIYDAVYGRIRSKVELEVR 382
 QY 361 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTIKPVG 420
 DB 383 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTIKPVG 442
 QY 421 FKOSLIQVQTFDDCAQQAAPNSHRNNGNTEFEGVRCRCPGMLGSCCESEEDYR 480
 DB 443 FKOSLIQVQTFDDCAQQAAPNSHRNNGNTEFEGVRCRCPGMLGSCCESEEDYR 502
 QY 481 SQODECSREGOPVCSQREGELCGQCCHSSDFKITGKICECDDEFCVYRKGEMSGHG 540
 DB 503 SQODECSREGOPVCSQREGELCGQCCHSSDFKITGKICECDDEFCVYRKGEMSGHG 562
 QY 541 QCSGDCGLCDSDMTGYCNCCTTRTDTTCMSSNGILCSGRGKCEGSCVCVCIQPGSYGDTCEK 600
 DB 563 QCSGDCGLCDSDMTGYCNCCTTRTDTTCMSSNGILCSGRGKCEGSCVCVCIQPGSYGDTCEK 622
 QY 601 CPTCPDACPCKKCEVCEKRRDGLHDENCNRYCRDELESYKELMDTGDAVNCYKNE 660
 DB 623 CPTCPDACPCKKCEVCEKRRDGLHDENCNRYCRDELESYKELMDTGDAVNCYKNE 682
 QY 661 DDCVVRFOYEDSSGKSILVVEEPECPKGPDLVLLSVGAAILLIGLAALLIMKLLLT 720
 DB 683 DDCVVRFOYEDSSGKSILVVEEPECPKGPDLVLLSVGAAILLIGLAALLIMKLLLT 742
 QY 721 IHRKEFAKFEERARAKMDTANNPLKKEATSTFTNITTYRGT 762
 DB 743 IHRKEFAKFEERARAKMDTANNPLKKEATSTFTNITTYRGT 784

RESULT 5
 Q9TUN5 PRELIMINARY; PRT; 784 AA.

AC Q9TUN5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE GLYCOPROTEIN IITA.
 GN GPIITA.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=PLATELET;
 RA Tao J., Parrilla R.;
 RT "Molecular cloning of pig platelet GPIIIa cDNA."
 RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
 CC 1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC 1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC 1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: AF170527; AAD51953.1;
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001169; Integrin_B.
 DR InterPro: IPR002369; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; vWFA.
 DR Pfam: PF00362; Integrin_B; 1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; vWA; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 3.
 DR Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix;
 KW Glycoprotein; Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 784 AA; 86399 MW; EDED5651EE288D5 CRC64;

Query Match 93.7%; Score 3894; DB 6; Length 784;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 706; Conservative 35; Mismatches 21; Indels 0; Gaps 0;

QY 1 GPNICITRGVSSCCOCLAVSPMCAMCSDDEALPLGSPRCDLKENILKNCAPESIEFPVSE 60
 DB 23 GPNICATRGVSSCCOCLAVSPMCAMCSDDEALPLGTPRCNLEENLNHCNPKIEFPISE 82
 QY 61 ARVLEDRPLSDKSGSDSSQTVSPQRIALRLPDDSKNFSIOVQVEDYPVDIYYLMDL 120
 DB 83 ARILEARPLSDKSGSDSSQTVSPQRIALRLPDDSKNFSIOVQVEDYPVDIYYLMDL 142
 QY 121 SYSKMDLMSIONLGTATLQATOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENFCYDM 180
 DB 143 SFSKMDLESIONLGTATLQATOMRKLTSNLRIGGAFYDKVSPMYTISPPEALENFCYDM 202
 QY 181 KTTCLPMFGYKHHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRN 240
 DB 203 KTTCLPMFGYKHHVLTLDQVTRFNEEYKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRN 262
 QY 241 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHYSASTTMDYPSLGLMTEKLS 300
 DB 263 DASHLVFTTDAKTHIALDGRLAGIVQPNDCQCHVSDNHYSASTTMDYPSLGLMTEKLS 322
 QY 301 QKNINLIFAVTVENYVNLQVNSSELIPGTGVLSMDSNVQLIYDAVYGRIRSKVELEVR 360
 DB 323 QKNINLIFAVTVENYVNLQVNSSELIPGTGVLSMDSNVQLIYDAVYGRIRSKVELEVR 382
 QY 361 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTIKPVG 420
 DB 383 DLPEELSLSFNATCLNNEVILPGKSCMGKLTIGTVSFSIEBAKVGCPQEKESFTIKPVG 442
 QY 421 FKOSLIQVQTFDDCAQQAAPNSHRNNGNTEFEGVRCRCPGMLGSCCESEEDYR 480
 DB 443 FKOSLIQVQTFDDCAQQAAPNSHRNNGNTEFEGVRCRCPGMLGSCCESEEDYR 502
 QY 481 SQODECSREGOPVCSQREGELCGQCCHSSDFKITGKICECDDEFCVYRKGEMSGHG 540
 DB 503 SQODECSREGOPVCSQREGELCGQCCHSSDFKITGKICECDDEFCVYRKGEMSGHG 562


```

QY 541 QCSGCDLSDSWTGYCNCRTTDTQMSNGLLSGRGKCEGCVCIQPSYGDTCER 600
DB 563 QCSGCDLSDSWTGYCNCRTTDTQMSNGLLSGRGKCEGCVCIQPSYGDTCER 622
QY 601 CPTCPDACCFFKCCVCKKFFDRCALHDEMTCNRYCDELESVELKDTCKDANCYKNE 660
DB 623 CPTCPDACCFFKCCVCKKFFDRCALHDEMTCNRYCDELESVELKDTCKDANCYKNE 682
QY 661 DDCVFRQYEDSSGKSTLYVEEPCPKGPDILVLLSMAIILIGLAALLIKLLIT 720
DB 683 DDCVFRQYEDSSGKSTLYVEEPCPKGPDILVLLSMAIILIGLAALLIKLLIT 742
QY 721 HDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTYRG 762
DB 743 HDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTYRG 784

RESULT 6
ID Q90W16 PRELIMINARY; PRT; 723 AA.
AC Q90W16;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DB BETA 3 INTEGRIN, GPIIA.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93290675; PubMed=8512576;
RA Cleutiaux A.M., Rosa J.P., Letourneur F., Poncez M., Rifat S.;
RT "A comparative analysis of cDNA-derived sequences for rat and mouse
RT beta 3 integrins (GPIIA) with their human counterpart.";
RL Biochem. Biophys. Res. Commun. 193:771-778(1993).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR Interpro: IPR002086; Aldehyde_dehydr.
DR Interpro: IPR000561; EGF-like.
DR Interpro: IPR002369; Integrin_B.
DR Interpro: IPR001169; Integrin_delta-C.
DR Interpro: IPR002035; VWA.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRIN.
DR PRODOM: PD001811; Integrin_B_1.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00187; VWA; 1.
DR SMART: SM00327; INB; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS00186; EGF_2; UNKNOWN_1.
DR PROSITE: PS00243; INTEGRIN_BETA_3.
KM Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
KW Integrin; Repeat; Transmembrane.
SQ SEQUENCE 723 AA; 80233 MW; B31D3127E6EC90F CRC64;

```

Query Match 86.88; Score 3606; DB 11; Length 723;
 Best Local Similarity 90.58; Pred. No. 3e-291;
 Matches 654; Conservative 42; Mismatches 27; Indels 0; Gaps 0;

```

QY 40 LKENLKNCAPESTIEFPSEARVLEDRPLSDKSGSSQYTVSPORIALRLRPDSKN 99
DB 1 LKNNLKKCKSPESIEFPSEARVLEDRPLSDKSGSSQYTVSPORIALRLRPDAKI 60
QY 100 FSIQVQVEDYVVDIYIMDLSTYKMDLMSIOMUGFTLAKOMKLSNLRIGGATVADK 159
DB 61 FSIQVQVEDYVVDIYIMDLSTYKMDLMSIOMUGFTLAKOMKLSNLRIGGATVADK 120
QY 160 PVSPTMYTISPPEALENPCYDKMTCTCLPMFGYKHVLTLLDQVTRFNEEVKQSVSRNDAP 219

```

```

DB 121 PVSPTMYTISPPEALENPCYDKMTCTCLPMFGYKHVLTLLDQVTRFNEEVKQSVSRNDAP 180
QY 220 EGGFPAIMQATYCDERKIGMRNDASHLFTYTDATHTIALDRLAGIYQVNDGQCHVSDN 279
DB 181 EGGFPAIMQATYCDERKIGMRNDASHLFTYTDATHTIALDRLAGIYQVNDGQCHVSDN 240
QY 280 HYSASTMDYPSLGLMTKLSOKNINLIFAVTENVNLYONSELIPCTYGVUSAMSSN 339
DB 241 HYSASTMDYPSLGLMTKLSOKNINLIFAVTENVNLYONSELIPCTYGVUSAMSSN 300
QY 340 VLQLLVDAVYGRKRSVLEVDLPDELSTSFNATCLNNEVPGKSCGKLGIDTVPSFI 399
DB 301 VLQLLVDAVYGRKRSVLEVDLPDELSTSFNATCLNNEVPGKSCGKLGIDTVPSFI 360
QY 400 EAKVRCGPOEKESFTIRPVGFKDSLIVQVTFDCCACQAQAEPNSHRCNNGNGTEECGV 459
DB 361 EAKVRCGPOEKESFTIRPVGFKDSLIVQVTFDCCACQAQAEPNSHRCNNGNGTEECGV 420
QY 460 CRCGPGWLGSCGCESEEDYRPSQDECSPREGQPYCPSRGCECLCGQCCHSSDFTKTK 519
DB 421 CRCGPGWLGSCGCESEEDYRPSQDECSPREGQPYCPSRGCECLCGQCCHSSDFTKTK 480
QY 520 YCECDDEFCVRYKGMCSGHGCGSCGDCDSDMTGYCNCRTTDTQMSNGLLSGRG 579
DB 481 YCECDDEFCVRYKGMCSGHGCGSCGDCDSDMTGYCNCRTTDTQMSNGLLSGRG 540
QY 580 KCEGSCVCIQPSYGDTCERKCPDACEFKKECKEDRGALHDEMTCNRYCDEI 639
DB 541 KCEGSCVCIQPSYGDTCERKCPDACEFKKECKEDRGALHDEMTCNRYCDEI 600
QY 640 ESKYELKDTCKDANCYKNEKEDYRPSQDECSPREGQPYCPSRGCECLCGQCCHSSDFTKTK 699
DB 601 ESKYELKDTCKDANCYKNEKEDYRPSQDECSPREGQPYCPSRGCECLCGQCCHSSDFTKTK 660
QY 700 VMAIILIGLAALLIKLLITIHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTY 759
DB 661 VMAIILIGLAALLIKLLITIHDRKEFAKFEERARAKMDTANNPLYKEATSTFTNTY 720
QY 760 RGT 762
DB 721 RGT 723

RESULT 7
ID Q92071 PRELIMINARY; PRT; 781 AA.
AC Q92071; Q90829;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DB INTEGRIN BETA3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WHITE LEGHORN;
RC MEDLINE=94164000; PubMed=8119143;
RA Miyama H., Cao X.P., Ross F.P., Chiba M., Teitelbaum S.L.;
RT "1,25-dihydroxyvitamin D3 transcriptionally activates the beta 3-
RT integrin subunit gene in avian osteoclast precursors.";
RL Endocrinology 134:1061-1066(1994).
RN [2]
RP SEQUENCE OF 1-26 FROM N.A.
RX TISSUE=BONE MARROW;
RC MEDLINE=94086557; PubMed=8262978;
RA Cao X., Ross F.P., Zhang L., MacDonald P.N., Chappel J.,
RA Teitelbaum S.L., Patrick F.;
RT "Cloning of the promoter for the avian integrin beta 3 subunit gene
RT and its regulation by 1,25-dihydroxyvitamin D3.";

```

RL J. Biol. Chem. 268:27371-27380(1993).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: X72378; CAAS1069.1; -
 DR EMBL: X75348; CAAS3095.1; -
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR pfam: PF00362; Integrin_B_1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD00181; Integrin_B_1.
 DR SMART: SM00001; EGF_Like_1.
 DR SMART: SM00187; INB_1.
 DR SMART: SM00423; PSI_1.
 DR SMART: SM00327; VWA_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR Cell adhesion: Cytoskeleton; Extracellular matrix; Glycoprotein;
 Integrin; Repeat; Transmembrane.
 KW
 SQ SEQUENCE 781 AA; 86115 MW; 7E991B605DF0CDBC CRC64;

Query Match 83.9%; Score 3484.5; DB 13; Length 781;
 Best Local Similarity 83.0%; Pred. No. 4.4e-281;
 Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;

QY 1 GNPICTRGVSSCOQCLAVSPMCWCSDEALPLGSPROCDLKNLTKNDCAPESEIFPSE 60
 DB 20 GSNICATRGVTSCKKCLAVSPICAMCSAENVAGSTPRCDLKNLQNGCGRFIFPPSS 79
 QY 61 ARYLERPISLSDKSGSDSSQVTVSPQIALRLRPDSKNFSIQVQVEDYPVDIYYLMDL 120
 DB 80 IYVERLPISLSDKSGSGST-TTQMSFQRIQNLRLPDSQMFVHVQVEDYPVDIYYLMDL 138
 QY 121 SYSKMDLWMSIONLGTATKATQMRKLTSLNRIGFGAFVDPKPSPVYATISPPALENPCDLM 180
 DB 139 SMSMDLKNIONLGTATKATQMRKLTSLNRIGFGAFVDPKPSPVYATISPPALENPCDLM 198
 QY 181 KTTCLPMFGYKHVLTLDQVTFEENEVKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 240
 DB 199 GEKCLPMFGYKHVLTLDQVTFEENEVKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRN 258
 QY 241 DASHLLVETQAKTHIALDGRAGIYQVNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLS 300
 DB 259 DASHLLVETQAKTHIALDGRAGIYQVNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLS 318
 QY 301 QKNINLIFAVTENYVNLKQNSSELIPGTVGLSDSSNNVQLIYDAVKIRSVLELVR 360
 DB 319 QKNINLIFAVTDYVGLKQNSSELIPGTVGLSDSSNNVQLIYDAVKIRSVLELVR 378
 QY 361 DLPEELSLSPNATCLNNEVIGLKSCKGLIGDYVFSIEAKVNGCDQEKESFTIPVVG 420
 DB 379 DLPEELSLSPNATCLNNEVIGLKSCKGLIGDYVFSIEAKVNGCDQEKESFTIPVVG 438
 QY 421 FKDSLIVQTPDCCACQAQAEPSNHRKNNNGNFTFEGCVCCGPGWLGSCQCESEEDYRP 480
 DB 439 FKDSLIVQTPDCCACQAQAEPSNHRKNNNGNFTFEGCVCCGPGWLGSCQCESEEDYRP 498
 QY 481 SQODECSPREGQVPCVSGEGELCGQCVCHSSDFGKINGKVCQEDDFGCVRRKKGKSGHG 540
 DB 499 SEQDNCSQPGQPLCSQREGEICQCVCHSSDFGKINGKVCQEDDFGCVRRKKGKSGHG 558
 QY 541 QCSGCDCLCSDMTGYCCNCTTRTDTCMSNGLICSGKCGKCGSCVCIQPGSYGDTCEK 600
 DB 559 QCSGCDCLCSDMTGYCCNCTTRTDTCMSNGLICSGKCGKCGSCVCIQPGSYGDTCEK 618
 QY 601 CPTCPDCTFKKCEVCECKFKEDRGALHDENVCNRYCRDEIESVKEKLDKGDVAVNCTYKNE 660

DB 619 CPTCPDCTFKKCEVCECKFKERGLVEQSCGVRDEIEYVEHGDGKDAVNCYKDE 678
 QY 661 DDCVRFQYEDSSGKSLIYVEEPCRGDIIIVLISWGCATILIGLALLKLLIT 720
 DB 679 NDCVRFQYEDSSGKSLIYVEEPCRGDIIIVLISWGCATILIGLALLKLLIT 738
 QY 721 IHDRKFAFEERARAKMDTANNPLYEATSTFTNITYRG 761
 DB 739 IHDRKFAFEERARAKMDTANNPLYEATSTFTNITYRG 779

RESULT 8
 Q9QW15 PRELIMINARY; PRT; 680 AA.
 AC Q9QW15.
 DT 01-MAY-2000 (Tremblrel_13, Created)
 DT 01-MAY-2000 (Tremblrel_13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel_19, Last annotation update)
 DE BETA 3 INTEGRIN, GPIIIB.
 OS Mus sp.
 OC Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93290675; PubMed-8512576;
 RA Cleat A.M., Rosa J.P., Letourneur F., Poncez M., Rifat S.;
 RT "A comparative analysis of cDNA-derived sequences for rat and mouse
 beta 3 integrins (GPIIIB) with their human counterpart."
 RL Biochem. Biophys. Res. Commun. 193:771-778(1993).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR002035; VWA.
 DR Pfam: PF00362; Integrin_B_1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD00181; Integrin_B_1.
 DR SMART: SM00001; EGF_Like_1.
 DR SMART: SM00187; INB_1.
 DR SMART: SM00327; VWA_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; INTEGRINB.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion: Cytoskeleton; Extracellular matrix; Glycoprotein;
 Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 680 AA; 75514 MW; 672DB738BD86003 CRC64;

Query Match 83.6%; Score 3472; DB 11; Length 680;
 Best Local Similarity 92.4%; Pred. No. 3.9e-280;
 Matches 628; Conservative 32; Mismatches 20; Indels 0; Gaps 0;

QY 79 QVTQVSPQIALRLRPDSKNFSIQVQVEDYPVDIYYLMDLSYSKMDLWMSIONLGTKL 138
 DB 1 QVTQVSPQIALRLRPDSKNFSIQVQVEDYPVDIYYLMDLSYSKMDLWMSIONLGTKL 60
 QY 139 ATQMRKLTSLNRIGFGAFVDPKPSPVYATISPPALENPCYDMKTTCLPMFGYKHVLTLD 198
 DB 61 ASQMRKLTSLNRIGFGAFVDPKPSPVYATISPPALENPCYDMKTTCLPMFGYKHVLTLD 120
 QY 199 QVTFEENEVKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRNDASHLLVETQAKTHIAL 258
 DB 121 QVTFEENEVKQSVSRNRDAPEGGFDAIMQATVCEDEKIGWRNDASHLLVETQAKTHIAL 180
 QY 259 DGRAGIYQVNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLSQKNINLIFAVTENYVNL 318
 DB 319 DGRAGIYQVNDGQCHVGSNNHYSASTTMDYPSLGLMTERKLSQKNINLIFAVTENYVNL 378

DB 181 DGRILAGIVLPNDGCHIGTNDHYSASTMDYSLGLMTEKLSOKNINLFAVTENVSLY 240
 QY 319 QNSELIPGTTVGLSDSSNNVQLLVDAVGRKRSKVELEVRDPELSLSPATCLNNE 378
 DB 241 QNSELIPGTTVGLSDSSNNVQLLVDAVGRKRSKVELEVRDPELSLSPATCLNNE 300
 QY 379 VIRGLKSCVGLKIGDVFSEIEAKVRCGPOEKESFTIKPVGFKDSLIVQVFPDCCACQ 438
 DB 301 VIRGLKSCVGLKIGDVFSEIEAKVRCGPOEKESFTIKPVGFKDSLIVQVFPDCCACQ 360
 QY 439 AQAEPSNRHRCNNGNFTFEGCVRCGPGWLGSCGCESEEDYRPSQODECSFREGQPVCSQR 498
 DB 361 AFAPSSPRCNCNNGTFFECVCVRCDOGWLSCMCSEEDYRPSQODECSFREGQPVCSQR 420
 QY 499 GECGCGCVCHSSDFGKITGKYCECDFSCVRYKKGKSCGCGSCGDCDCLDMDTGYCC 558
 DB 421 GECGCGCVCHSSDFGKITGKYCECDFSCVRYKKGKSCGCGSCGDCDCLDMDTGYCC 480
 QY 559 NCTRTDTGMSNGLCSGRKCEGSCVCIQPGSYGDTCEKCPDCTCFKKECCECK 618
 DB 481 NCTRTDTGMSNGLCSGRKCEGSCVCIQPGSYGDTCEKCPDCTCFKKECCECK 540
 QY 619 KEDRGALHDENTCNRYCRDEIESYKELKDTGKDAVNTYKNEDDCVYRFGYEDSSGSKI 678
 DB 541 KFNRTGHEENTCSRYCRDIEQYKELDTGKNAVNTYKNEDDCVYRFGYEDTSGRAV 600
 QY 679 LYVEEPCPCPGDILVLLSVMGAILLIGLALLIKLITTIHDKREFAKFEERARAK 738
 DB 601 LYVEEPCPCPGDILVLLSVMGAILLIGLALLIKLITTIHDKREFAKFEERARAK 660
 QY 739 WDTANNPYKCATSTFTNIT 758
 DB 661 WDTANNPYKCATSTFTNIT 680

RESULT 9
 007012 PRELIMINARY: PRT: 788 AA.
 ID 007012
 AC 007012
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE INTEGRIN BETA-3 SUBUNIT PRECURSOR.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 CC Xenopodidae; Xenopus.
 CC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94040374; PubMed=7693527;
 RA Ransom D.G., Hens M.D., Desimone D.W.;
 RT "Integrin expression in early amphibian embryos: cDNA cloning and
 RT characterization of Xenopus beta 1, beta 2, beta 3, and beta 6
 RT subunits."; Biol. 160:265-275(1993).
 RL Dev. Biol. 160:265-275(1993).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL, L13591: AAA17427.1; -;
 DR InterPro: IPR002086; A1dehydratase.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PF00362; Integrin_B.1.
 DR PRINTS: PRO1186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B.1.
 DR SMART: SM00181; EGF.1.
 DR SMART: SM00187; INB.1.
 DR SMART: SM00423; PSI.1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.

DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA_2.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 788 INTEGRIN BETA-3 SUBUNIT.
 SQ SEQUENCE 788 AA: 87490 MW: 80902510E24A6109 CRC64:

Query Match 78.5%; Score 3260.5; DB 13; Length 788;
 Best Local Similarity 75.9%; Pred. No. 1.9e-262;
 Matches 578; Conservative 89; Mismatches 94; Indels 1; Gaps 1.

QY 1 GPNICTTRGSSCCQCLAVSPMCANCSDEALPLGSPRCDLKENLKDNCAPESIFPVS 60
 DB 24 GASICATRGSSCCQCLAVSPMCANCSDEALPLGSPRCDLKENLKDNCAPESIFPVS 83
 QY 61 ARVLEDRPLSDGS-GDSQVQVTPORTALRLRBDSDKNSIOVROYEDYPVDTYIAMD 119
 DB 84 VLVAENRPLSVKSGSDREITQMSPOKIDFLRPDDTKVFMLOVROYEDYPVDTYIAMD 143
 QY 120 LSYMKDDLSIONLGTATQMRKLTSLRIGFAFYDKPVSPYMYTSPPEALENPGYD 179
 DB 144 LSYMKDDLIKOTLTSTLSERMRLTSLRIGFAFYDKPVSPYMYTSPPEALENPGYD 203
 QY 180 MTTCLPMFGYKHVLTLDQVTRFNEVYKQSVSRNRDAPESGFDAINQATYCDKIGWR 239
 DB 204 FTECHPMFGYKHVLTLEVRNFNEVQKQVSRNRDAPESGFDAINQATYCDKIGWR 263
 QY 240 NDASHLLVFTTQAKHIALDGRLAGIVOPNDGCHGNSNHYASSTNDYPSLGLMTEKL 299
 DB 264 NESHLVFTTQAKHIALDGRLAGIVOPNDGCHGNSNHYASSTNDYPSLGLMTEKL 323
 QY 300 SOKNINLFAVTENVVLYONTSELIPGTTVGLSDSSNNVQLLVDAVGRKRSKVELEV 359
 DB 324 SOKNINLFAVTENVVLYONTSELIPGTTVGLSDSSNNVQLLVDAVGRKRSKVELEV 383
 QY 360 RDLPEELSLSFNATCLNNEVYIPGLKSCMGKIGDYVSEIEAKVRCGPOEKESFTIKPV 419
 DB 384 RDLPEELSLSFNATCLNNEVYIPGLKSCMGKIGDYVSEIEAKVRCGPOEKESFTIKPV 443
 QY 420 GFQDSILVQVTFPCDCAQAOAPNSHRNNGNGTFFEGCVRCGPGWLGSCGCESEEDYR 479
 DB 444 GFQDSILVQVTFPCDCAQAOAPNSHRNNGNGTFFEGCVRCGPGWLGSCGCESEEDYR 503
 QY 480 PSQODECSFREGQPVCSQRDEICLGQCVCHSSDFGKITGKYCECDFSCVRYKKGKSCG 539
 DB 504 PSQODECSFREGQPVCSQRDEICLGQCVCHSSDFGKITGKYCECDFSCVRYKKGKSCG 563
 QY 540 GCGSCGDCDCLDMDTGYCCITRTDTGMSNGLCSGRKCEGSCVCIQPGSYGDTCE 599
 DB 564 GCGSCGDCDCLDMDTGYCCITRTDTGMSNGLCSGRKCEGSCVCIQPGSYGDTCE 623
 QY 600 KCPTCPDCTCFKKECCECKKFDGALHDENTCNRYCRDEIESYKELKDTGKDAVNTYK 659
 DB 624 KCPTCPDCTCFKKECCECKKFDGALHDENTCNRYCRDEIESYKELKDTGKDAVNTYK 683
 QY 660 EDDCVARFQYEDSSGSKILYVVEEPCPCPGDILVLLSVMGAILLIGLALLIKLIT 719
 DB 684 EDDCVARFQYEDSSGSKILYVVEEPCPCPGDILVLLSVMGAILLIGLALLIKLIT 743
 QY 720 TIHDKREFAKFEERARAKWDTANNPYKCATSTFTNITRG 761
 DB 744 TIHDKREFAKFEERARAKWDTANNPYKCATSTFTNITRG 785

RESULT 10
 09GK49 PRELIMINARY: PRT: 791 AA.
 ID 09GK49
 AC 09GK49
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE INTERLIN BETA-5 SUBUNIT (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NCBI_Taxid=9913;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP TISSUE=MAMMARY GLAND;
 RA Andersen M.H., Rasmussen J.T., Berglund L., Petersen T.E.;
 RT "Bovine beta-5 integrin subunit (fragment)."
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317198; AAG36594.1;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PR00362; Integrin_B.
 DR PRINTS: PRO1186; INTEGRIN_B.
 DR ProDom: PD00181; Integrin_B; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00001; EGF_like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN.1.
 DR NON_TER 1
 FT SEQUENCE 791 AA; 87144 MW; 3FB45E62374169BC CRC64;

Query Match 55.7%; Score 2315.5; DB 6; Length 791;
 Best Local Similarity 54.8%; Pred. No. 8.4e-184;
 Matches 426; Conservative 122; Mismatches 211; Indels 19; Gaps 9;

QY 1 GNICTRTGSSCOCCIAVSPMCACSD---ALPLGSPRODLKKNLKDNCAPSIIEP 57
 DB 15 GNICTSGATSCCECLLHPGACWCFEDGSLSVSRCDLKNLRNCCGVE-FESP 73
 QY 58 VSEANLEDRPLSDKSGDS--SQVTSVPORIALRLPDDSKNFISIQROYEDYVVDIY 116
 DB 74 ASSTQVLSLPLSSKSGSPAGSDVQLPQEVTVTLRPGDTRATQLOVROVEDYVVDIY 133
 QY 117 LMDLSYKDDLMSTQNGTKIATQMKRLTSLNRIGFAYVDKVPYMYISPEALENP 176
 DB 134 LMDLSLMSKDDLDNIRSLGTIKLAEMRKLTNFRIGFSFVDKNTSPSY-TAPRYQNP 192
 QY 177 C--YDMKTTCLPMFGYKRVLTLDQVTRFNEEVKKQSSRRRDAPEGFDLIMATYCD 234
 DB 193 CIGYLPNCVPSFGFRHLPLTDVDSFNEEVKQSSRRRDAPEGFDLIMATYCD 252
 QY 235 KIGWRNDASHLLVFTTDAKTHIALDGLAGIVQPNDCQCHVGSNNHYSASTTMDYPSLG 294
 DB 253 KIGWRKDALHLVFTTDOVPHIALDGLGIVQPHDQCHLNEANETASQMDYPSIAL 312
 QY 295 MTEKLSQNTLILFAYENVNLYQNTSELIPGTVGVLSMDSNVQLIIVDAYGKTRSK 354
 DB 313 LGKILAEKNINILFAYVKNHMYLKNFTALIPGTVVEIHLHDSKNIILIIINAVYSIRSK 372
 QY 355 VELEVRDLPBELSLSPNATCLNNEVIPGKSCMGLKIGDTPYSPFIEAKVGCQPEK 413
 DB 373 VELVMPQPEDLNLFATATCDGVSYEQKRCBELKIGDTRASLRVSEARSCPKHVEHT 432
 QY 414 FTIKPVGKDSLIVQVTFDDCAQAOAEPNSHRNNGNGTFFEGVRCRGPWGLSQCEC 473
 DB 433 LALRPVAFRDSLEGVYVYNCRCGAGLEPDSARCS-NGIYVAGLECPNGYLGTCCEC 491
 QY 474 SEELYRPSQDECPREGQPVYCSRGKCLGQCYCHSDGKTKTGKCEBDPSCVARYK 533
 DB 492 QEGESQGYQMLCEARAGKPLCSGRQSCNQCSCFSESEKTIYGSFCEDNFCARANK 551
 QY 534 EMCSGHQSCGDLCDSDMTGYCNCCTTRDPCMSNGILGSRGCEGSCVCVCIOPGS 593
 DB 552 VLSGHECHGECEKCHAGYIDNCNSTDSTOCARDGHICSDRGHCVCQCCCTEPEGA 611
 QY 594 YDTCERCEPTCPDACEKFEKCECKFRGALHDENTONRCYRDE-IESVELMDTGDA 652
 DB 612 FGETCEKPTCPDACEKFEKCECKFRGALHDENTONRCYRDE-IESVELMDTGDA 671
 QY 653 VNCYKKNEDDCVRFQYVEDSSGKSLVVEPECPKPDILVLLVLSYMAIILIGLAAL 712
 DB 672 VLSGHECHGECEKCHAGYIDNCNSTDSTOCARDGHICSDRGHCVCQCCCTEPEGA 731
 QY 713 LMKLITITIHDRKEFAKFEERARAKMDTANNPYKATSTFT-----NITRYGT 762
 DB 722 VIMKLVTIITHDRKEFAKFEERARAKMDTANNPYKATSTFT-----NITRYGT 789

RESULT 11
 ID O9GLP0 PRELIMINARY; PRT; 798 AA.
 AC O9GLP0;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE INTERLIN BETA-1 SUBUNIT.
 GN CD29.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_Taxid=9623;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP MEDLINE-20426040; Pubmed-10972224;
 RA Jimenez-Marin A., Garrido J.J., de Andres-Cara D.F., Moreta L.,
 RA Barbancho M.J., Llanes D.;
 RT "Molecular cloning and characterization of the pig homologue to human
 RT CD29, the integrin beta1 subunit."
 RL Transplantation 70:649-655(2000).
 DR EMBL; AF192528; AAG16767.1;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PR00362; Integrin_B.
 DR ProDom: PD00181; Integrin_B; 1.
 DR PRINTS: PRO1186; INTEGRIN_B.
 DR ProDom: PD00181; Integrin_B; 1.
 DR SMART: SM00001; EGF_like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN.1.
 DR SHOUENCE 798 AA; 88265 MW; 80773E95E38BA90 CRC64;

Query Match 44.1%; Score 1833; DB 6; Length 798;
 Best Local Similarity 44.2%; Pred. No. 1.2e-143;
 Matches 345; Conservative 140; Mismatches 267; Indels 28; Gaps 13;

QY 3 NICTTGVSQCCOCCIAVSPMCACSD---EALPLGSPRODLKKNLKDNCAPSIIEP 57
 DB 25 NIKLAKNAKSGCEGCIQAGPNCWCNTSTLQEGMP-SARCDLELRKKGCHPDIEP 83
 QY 58 VSEARVLEDRPLSDKSGDSQ---VYQVSPORIALRLPDDSKNFISIQROYEDYV 112
 DB 84 RGSKNIKKKNVYTRNSKGAELKQPEIDIQIPQOLVLOLBSGEQOTFLTKKRAEDYPI 143
 QY 113 DIYIMLDSYKDDLMSTQNGTKIATQMKRLTSLNRIGFAYVDKVPYMYISPEEA 172
 DB 144 DIYIMLDSYKDDLMSTQNGTKIATQMKRLTSLNRIGFAYVDKVPYMYISPEEA 202
 QY 173 LENPCDMKTTCLPMFGYKRVLTLDQVTRFNEEVKKQSSRRRDAPEGFDLIMATYCD 232
 DB 203 LKRPC-TSQNCTSPFSYKYNVSLTDKGEVVELYKORISLNDSPGCGDPAIMQVAVC 261
 QY 233 DEKIGWRNDASHLLVFTTDAKTHIALDGLAGIVQPNDCQCHVGSNNHYSASTTMDYPSL 292

DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE BETA INTEGRIN SUBUNIT
 OS Biomphalaria glabrata (Bloodfluke planorb).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 CC Planorbidae; Biomphalaria.
 NCBI_TaxID=6526;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99173885; PubMed=10072774;
 RA Davids B.J., Wu X.J., Yoshino T.P.;
 RT "Cloning of a beta integrin subunit cDNA from an embryonic cell line
 RT derived from the freshwater mollusc, Biomphalaria glabrata."
 RL Gene 228:213-223(1999).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: AF060203; AAC67503.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR002049; Laminin_EGF.
 DR Pfam: PR00362; Integrin_B_1.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B_1.
 DR SMART: SM00001; EGF_like_1.
 DR SMART: SM00187; INB_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 788 AA: 87632 MW: 3FE3DFAFE0848163 CRC64:

Query Match 39.9%; Score 1658; DB 5; Length 788;
 Best Local Similarity 42.7%; Pred. No. 4.3e-129;
 Matches 329; Conservative 136; Mismatches 267; Indels 38; Gaps 20;
 DB 12 SCQCCLAASPMCAWCSDEALPLGSP-RCDIKENLLKD-NCAPESIEFVSEKRVLEDPRL 69
 DB 32 SCASCIATSKCECAWCVSYEDENLRCDTHENHRLGALYCPGDIQFPTDEVEKLKNO-- 89
 QY 70 SDRGSGDSQ-VTQVSPQRIALRLRPDSSKNFSIQVQVEDPVYIYIMLSYSMKDDL 128
 DB 90 -DVGGDNDPDAVQVQPKVRIKIRPNKPEVKKLFRGAENYVDLLIFMLDSMEDDK 148
 QY 129 WSIQWLGRKLATQMKLTSNLIGGAFVYKPSYMYTISPEALENFCYDMKT---TC 184
 DB 149 EKLALLGKRIAEQMSAIRKRNRLGFGSPVDVVSPPY-STVPOKLMKC--KITNGEPC 204
 QY 185 LPMFGYKVLTLTDQVTFENNEVKQSVSRNRDAPEGGEDALMATVQDEKIGRRNDASH 244
 DB 205 EAPYFKKQSLIDLETTFESQKVFARSGVLDAPEGGFDAIMQVACEDIEDIGRRPSRR 264
 QY 245 LLVPTTKTHIALDGRAGIIVQPNQGCCHVSDNHYSASTTMDYPSLIGLTKELSKONI 304
 DB 265 MLVSTDAGFHAGDGKIGIIVTPNDGCHL-KNNLYSESSNLDYPSVSOJANKIKRSV 323
 QY 305 NLITAVTENNVNIQNSYSELPGTTVGVLSMDSNVQIOLYDAGKIRSKVELEVRDLPE 364
 DB 324 SVITAVVDLPDIYEKLSKYESSITGLANDSSNIYLLIDENYKTKITSKVTLAADIGE 383
 QY 365 ELISFNATCLINNEVITGLKSCMGLKIGDYVSFSIEAKVRGCPQEK---KSFTIKPVGF 421
 DB 384 NITYDFSRCHRGGEIK-TNCGSLKIGQSVFPAEVLNLRCPKDRKKWLKERSIRPLGY 442
 QY 422 KDSLIIVGTFFCCACQ--AQAEFNSHRCNNGNGTFCGVCRCGPGWLGSCCECSEEDYR 479
 DB 443 QEKLTVELEMICCEOCENAEENEILNSDKCSNGNGTFCGCKSCGHPRGYGFCECKADDLT 502

QY 480 P8QD-DEGSPREGQVYCSORGECLGOCVCH--SSDFPK-ITGKYCEDDFSVRYKREM 535
 DB 503 SKDSIKQCAPAPATLPCSRGSCVCGECLCNRSSDSQTSRGPCEDDVSCNOFNQI 562
 QY 536 CSG-HGQSCGDCICDSDMTGYNCCTTRIDTQSSNGLLCSGRKCEGSCVCIQGS 593
 DB 563 CGSPERKCNKCGCKCKLIGNCTACELSMEKCTTDDGLVGNHGNCTCKKVC--ENQ 620
 QY 594 Y-GDTCEKCPCTPDACITRKCEVCKKFDGALHDENTCNRYCDEI--ESVRELDTG 649
 DB 621 YTFKCEQCPACPDKCLEYTPCVCQKAFKGLQER-CLRECKMEIFHDKIRE---G 675
 QY 650 KQAVCTYKNEDDQVVRQYEDSSGKILVVEPECKRPDILVLLSMGAILLGL 709
 DB 676 HGICLCPKIDDECAVITYEYDRGKVIAQITKYCPDGLNVALVGGVGVAVGL 735
 QY 710 AALLTWKLITITIDRKEFAKFEERAKAMDTPANNPLKRETSFTNTITY 759
 DB 726 FULLIMKLTLFIDHREFAKFEKRONKMDTNGENPIYKATSTFKNPTY 785

RESULT 14
 ID Q9NAS7 PRELIMINARY; PTI; 837 AA.
 AC Q9NAS7;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE INTEGRIN BETA SUBUNIT PRECURSOR.
 GN BINT.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 CC Anopheles.
 NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2131701; PubMed=11437913;
 RA Machairakti V., Lycett G., Blass C., Louis C.;
 RT "Beta-Integrin of Anopheles gambiae: mRNA cloning and analysis of
 RT structure and expression."
 RL Insect Mol. Biol. 10:217-223(2001).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: AJ292755; CAC00630.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; VWFA.
 DR Pfam: PF00362; Integrin_B_1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B_1.
 DR SMART: SM00187; INB_1.
 DR SMART: SM00423; PSI_1.
 DR SMART: SM00327; VWFA_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Signal; Transmembrane.
 FT SIGNAL 1
 FT CHAIN 27 837 POTENTIAL.
 FT SEQUENCE 837 AA: 92791 MW: 30F1F3396924D78 CRC64;

Query Match 38.7%; Score 1609; DB 5; Length 837;
 Best Local Similarity 39.5%; Pred. No. 5.6e-125;
 Matches 323; Conservative 138; Mismatches 284; Indels 72; Gaps 19;

OY 4 ICTTBVSSGQOOLAVNSPIMCANSDBALP-LCSPRCDLKENLKNCAPESTIEFPVSFAR 62
 Db 33 LITTCPEKKTCSOCIOTFT--NCRCMT--MNFTHPRC---HGQIEKYCEPEYTVDPSPNFG 85
 OY 63 VLEDRPL-----SDKSGSDGS-----QVTVGS 84
 Db 86 LVQGRRLTTPSRRLVLEGQSERSESYSSSHYOGSSSSSSSSSSSPQSSYSESGAGSIYIS 144
 OY 85 POHIALRLAPDSSKNFSIOVROVEDYIPVDIYTLMDLSTSMKDDLWISIONLGTKLTATOMK 144
 Db 146 POWVSATKLNLNEAFRENVNVAQAEYPPVDLYIYMLDLSKSMEDDKITLSTGLADLASEMRK 205
 OY 145 LTSNLTGGAFAVDKRVSPYNYITSPPEALENCYDMKTKCLPMCEYKXVLLVLTQVYREN 204
 Db 206 IITSNFLGCGSYDVKLAMYV-STYPKNLREPC----PGCAVPYGHMLMLSTDANLFS 266
 OY 205 EBYVKOSVARNDAPEGFDAIMQATVYCDKIGMRNDASHLVFTTDAKTHIALDGLAG 266
 Db 261 QEVORANVSGNLDAPEGGFDAIMQATVYCDKIGMRNDASHLVFTTDAKTHIALDGLAG 322
 OY 265 IYQPNQCGHVSNDHYSASTMTDYPISGLMTERKISQKNINLIFAVTENVNLVONYSEL 324
 Db 321 VTPPDNGECHLDHNRKRYHSTTQDPYPSIOJNLKWKQNAIVYLAIVAEELSYEQLSRL 380
 OY 325 IGTGTGVLSMDSSVNLQIYAVYKIKISKEYLEVRDLPEELSLSFNATCLN-NEVYPL 388
 Db 381 VEGSSAARKLSNDSNINVSILYRDQYIKKISSYEMKNDFTDNIYDKVYSRCNTNGALQOT 440
 OY 384 KSCMGKLTGDIYVSFSIEAKVRGC---POEKESFTIKFVGKDSLIVQVTFDCDCAQO 440
 Db 441 NNCCEGLKVGDDVYTFEAHITLLKCPDPDPMQOVQIDYVPGINESLTVDIEMLSCPEHP 500
 OY 441 AEP-----NSHRCNNGNFTFEGVCYCGGMYLCSQCEGS--BEDYRPSQODECSFREQPV 494
 Db 501 SPPETREARDECSNA-GTYKCGICGCDGTTHGQRCESAMSSLLPBGWADCRMSNASEE 559
 OY 495 CSGRGECLGQGCVC--HSSDFKLTIKKCYCECDDFSCVAYKEMCSG--HGQCSGDCICD 550
 Db 560 CSGRGQCVQCVYCBRRPNPDELIDGRVCECDNFSQCRPGGLLCSGPDHGVCYGGQCECR 619
 OY 551 SPMVGYENCNTRITRTYCH--SSNGLLCSRGKCEGSGVC--TOPGSY-GDTCEKCPYCPD 606
 Db 620 EGMVTPADDCRASNETCMPPGGGLSCGHGTCCEGTCRCYTVEDGRRYTGRCYCEKCPYCG 679
 OY 607 ACTPFKECVCECKKPFRLALHDENFCNRCR--DEIESYKELDQGRKAVNCTYKNEDCV 664
 Db 680 RCNEKRKHVCCOQYKTGFLAANECATNCTLPVPIVEKATYIDERNMNCITFEDEDCR 739
 OY 665 VAFQYEDSSGKSLIYVEEPECEKGPDIIVLVLVSWGALITLIGLAILMLKLTITTHDR 724
 Db 740 PEFYSVNDSDQKVVYTAQENRECPKPKFMLOIYALVAVVYLIMAVLLMKVLTSTIHR 799
 OY 725 KEFAKFEERAKAKWDIANNPLVEATSTETNTIYRG 761
 Db 800 REFARFEKERMAKWDGEMPIYKQATVTTFKPNYYAG 836
 RESULT 15
 P92163 PRELIMINARY; PRT; 783 AA.
 ID P92163
 AC P92163;
 DT 01-MAY-1997 (TREMBLREL_03, Created)
 DT 01-MAY-1997 (TREMBLREL_03, Last sequence update)
 DT 01-DEC-2001 (TREMBLREL_19, Last annotation update)
 DE INEGRIN BETA G SUBUNIT.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echiniozoa;
 OC Echinioidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotidae;
 CX NCBI_TaxId=7668;
 RP [1]
 SEQUENCE FROM N.A.

RA Marsden J., Burke R. D. ;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1 PDM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL: U77584; AAB39739.1; -.
 DR EMBL: U77587; AAB39741.1; -.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002035; vWFA.
 DR Pfam: PF00362; Integrin_B; 1.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00327; vWA; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA; 2.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Transmembrane.
 SQ SEQUENCE 783 AA; 85530 MW; BB045C6F6D88FBB2 CRC64;

[illegible]

Tue May 21 14:53:25 2002

us-09-673-302a-1.rsp

Page 12

Db 614 NASSYRGALQDDEPTSGGCSRNEECVQKAFGTG- -SKAPQDK-CEPPVIMVDALEI 670

Qy 647 DTGDANVCTKNNDDCVPROXYEDSSGKSLIIVYEPCPGPDLIVLLSVMGAIL 706

Db 671 PTGSR-RCLEADEDGCSLIIFYTKASNMLALILVYQKEKCFEYVDIMHIIIGIVGII 728

Qy 707 IGLAALLIWKLLITIIHDKRFAKKEERAKRKYDTANPLIKYKATSEFTTITR 760

Db 729 VGLAILIWKRLIIVYDQSRERASTEKERAGIHGMQGNENPLIKPSTISFKNPYQ 782

Search completed: May 19, 2002, 12:26:43
Job time: 6465 sec